# İnsayder Təhdidlərini Aşkarlamaq üçün Süni İntellekt Metodlarının Analizi

Fərqanə Abdullayeva
İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan
*a_farqana@mail.ru*

***Xülasə*— Kibertəhlükəsizlik sahəsində ən təhlükəli hücumlar çox vaxt xarici hücumçulardan deyil, sistem daxilində etibarlı istifadəçilərdən qaynaqlanan insayder təhdidləridir. Bu hücumlar istifadəçilərin sistem resurslarına icazəli çıxış imkanlarına malik olması səbəbindən təhlükəsizlik mexanizmlərindən asanlıqla yan keçə bilir və aşkarlanması çətin olur. Məqalədə insayder təhdidlərinin aşkarlanması və qarşısının alınması sahəsində mövcud yanaşmalar təhlil edilmişdir. Xüsusilə istifadəçi davranışının analizi, əlamətlərin seçilməsi, maşın təlimi və dərin təlim metodlarının tətbiqi geniş şəkildə araşdırılmışdır, insayder hücumlarının növləri və onların qarşıdurma mexanizmləri təhlil edilmişdir. İşdə həmçinin mövcud problemlər, açıq tədqiqat məsələləri və gələcək inkişaf perspektivləri müəyyən edilərək təkliflər verilmişdir.**

***Açar sözlər*— *insayder təhdidlərinin aşkarlanması; loq faylların analizi: istifadəçi və obyekt davranışının analizi; kölgə süni intellekti; maşın təlimi***

## I. GİRİŞ

Müasir informasiya texnologiyalarının sürətli inkişafı və təşkilatların rəqəmsal infrastrukturdan geniş istifadəsi kiber təhlükəsizlik sahəsində yeni və daha mürəkkəb təhdidlərin yaranmasına səbəb olmuşdur. Bu təhdidlər yalnız kənar mənbələrdən deyil, həmçinin təşkilatın daxilindən insayderlər tərəfindən qaynaqlanır. İnsayderlər təşkilatın daxilindəki digər əməkdaşlarla müqayisədə daha çox etibar edilən və imtiyazları yüksək olan şəxslərdir [1]. İnsayderlər həssas məlumatları istismar etməklə və ya ifşa etməklə təşkilatlara ciddi təhdid yarada bilərlər. Təşkilatın daxilindəki insayderlər tərəfindən törədilən zərərli əməliyyatlar insayder təhdidi hesab olunur.

İnsayder təhdidi – hazırkı və ya keçmiş əməkdaş, podratçı və ya digər biznes tərəfdaşı tərəfindən təşkilatın şəbəkəsinə, sisteminə və ya məlumatlarına qanuni giriş hüququna malik olub (və ya olmuş) şəxsin həmin giriş hüququndan qəsdən (və ya bilmədən) sui-istifadə etməsi və ya səlahiyyət həddini aşması nəticəsində təşkilatın məlumatlarının və ya informasiya sistemlərinin məxfiliyinə, bütövlüyünə və ya əlçatanlığına mənfi təsir göstərməsidir [2]. İnsayderlərin təşkilatın informasiya sisteminə giriş səlahiyyətinin olması onların aşkarlanmasını xarici təhdidlərlə müqayisədə olduqca çətinləşdirir.

Təşkilatın ən mühüm aktivləri onun maliyyə məlumatları, kommersiya sirləri, intellektual mülkiyyəti və nüfuzudur. Bu aktivlərin hər hansı biri işçi tərəfindən bilərəkdən və ya bilmədən həyata keçirilən məlumat sızması nəticəsində ifşa olunduqda ictimai etibarın pozulması ilə təşkilatın reputasiyasına ciddi zərər dəyə bilər. Eyni zamanda, maliyyə xarakterli məxfi məlumatların sızması təşkilatın strategiyaları haqqında həssas detalları üzə çıxara bilər.

Avropanın insayder risklərinin idarə edilməsi üzrə aparıcı Signpost Six təşkilatının hesabatında qeyd edilir ki, 2025-ci ildə dünya təşkilatlarının 76%-i insayder insidentlərinin təsirinə məruz qalmışdır [3]. Ponemon İnstitutunun 2016-cı ildə dərc etdirdiyi "Cost of Insider Threats Global Report, 2016" adlı hesabatında verilənlərin sızması ilə bağlı 874 insident qeydə alınmışdır [4]. Bu insidentlərin 568-i insayderlər tərəfindən törədilmişdir. Ponemon İnstitutu tərəfindən 2022-ci ildə dərc edilmiş "Cost of Insider Threats Global Report, 2022" hesabatında qeyd edilmişdir ki, 2020-ci illə müqayisədə insayder təhdidi insidentləri 2022-ci ildə 44% artmış, hər bir insidentin vurduğu ziyanın dəyəri isə 15.38 milyon dollara çatmışdır [5]. Ponemon İnstitutunun "2025 Cost of Insider Risks Global Report" hesabatına görə, 2018-2024-cü illər arasında insayder təhlükə hadisələrinin ümumi orta dəyəri 109%-dən çox artmışdır [6]. Dünya üzrə bir təşkilat üçün insayderlə əlaqəli hadisələrin vurduğu orta zərərin həcmi 2025-ci ildə 17,4 milyon ABŞ dolları, 2026-cı ildə isə 19,5 milyon ABŞ dolları təşkil etmişdir. Bu göstərici 2025-ci illə müqayisədə 2026-cı ildə 12% artmışdır [7]. Məlumat sızmalarının 30%-i daxili aktorlarla əlaqələndirilir və təşkilatların 76%-i insayder hücumlarının daha tez-tez baş verdiyini bildirir. Təhdid yalnız zərərli əməkdaşların məxfi məlumatları oğurlaması ilə məhdudlaşmır, insayder hadisələrinin 55%-i səhlənkarlıq və insan səhvlərindən qaynaqlanır. Bu kimi müxtəlif və geniş təhlükə ssenarilərinin mövcudluğu bu tip potensial risklərin qarşısını almaq üçün yeni texnologiyaların işlənməsini aktuallaşdırır.

Zərərli insayderlərin qarşısını almaq üçün təşkilatlarda təhdidləri həyata keçirməzdən əvvəl onları aşkarlaya və qarşısını ala bilən insayder təhdidlərinin aşkarlanması sistemləri mövcud olmalıdır. Lakin insayder təhdidləri sahəsi hələ də kifayət qədər araşdırılmamışdır. Bundan əlavə, istifadə olunan aşkarlama mexanizmləri və yanaşmaları, eləcə də mövcud həllərin məhdudiyyətləri tam şəkildə öyrənilməmişdir. Buna görə də insayder təhdidlərinin aşkarlanması sahəsində mövcud yanaşmaların icmalının aparılması zəruri hesab olunur.

Təqdim olunan məqalədə insayder təhdidlərinin növləri və real insayder hücumu faktları araşdırılır, onların aşkarlanması üzrə mövcud yanaşmalar təhlil edilir. Bundan əlavə, kölgə süni intellektinin insayder təhdidlərinin yaranmasında rolu, istifadəçi davranışının analizi və bu sahədə tətbiq olunan məşhur hücum nümunələri təqdim olunur. Məqalədə insayder təhdidlərinə qarşı mübarizə strategiyaları və onların aşkarlanmasında süni intellekt metodlarının tətbiqi, kibertəhlükəsizlik sahəsində aparıcı təşkilatların mövcud alətləri tədqiq edilir.

DOI: 10.25045/forensic.gov.2026.06

## II. İNSAYDER TƏHDİDİ AKTORLARI

İnsayder hücumları təhdid aktorları tərəfindən törədilir. İnsayder təhdidləri aktorlarına aşağıdakılar aiddir:

İmtiyazlı istifadəçilər (privileged users). Yüksək giriş səlahiyyətləri olan istifadəçilərdir. Məsələn, administratorlar.

Sıravi işçilər (regular employees). Sıravi işçilərin imtiyazlı istifadəçilərlə müqayisədə imkanları məhduddur, lakin onlar yenə də təşkilata zərər verə bilərlər. Məsələn, onlar təsadüfən korporativ məlumatlardan sui-istifadə edə, icazəsiz tətbiqlər quraşdıra, məxfi e-poçtları səhv ünvana göndərə və ya sosial mühəndislik hücumunun qurbanı ola bilərlər. Lakin bütün işçilər təşkilatı təsadüfən təhlükəyə atmırlar. Zərərli işçilər çox vaxt maddi qazanc, qisas və ya ideoloji səbəblər niyyəti ilə hərəkət edirlər.

Üçüncü tərəflər, kontraktorlar (third parties). İT sistemə və ya məlumatlara çıxışı olan təchizatçılar, subpodratçılar, biznes tərəfdaşları və təchizat zənciri qurumlarıdır. Bu şəxslər əməkdaşlıq etdikləri təşkilatın kiber təhlükəsizlik qaydalarına əməl etmir və ya bilərəkdən zərərli hərəkətlər edirlər. Digər tərəfdən hakerlər birbaşa təşkilatın əsas sisteminə hücum etmək əvəzinə, zəif qorunan bir tərəfdaş şirkəti (vendoru) hədəfləyir, oradan keçərək təşkilatın qorunan sisteminə daxil olurlar.

Keçmiş əməkdaşlar (former employees). Təşkilatdan ayrılmış (işdən çıxmış və ya öz istəyi ilə tərk etmiş) şəxslərdir, əvvəlki giriş hüquqlarından, biliklərindən və ya əldə etdikləri məlumatlardan istifadə edərək sistemə zərər verir və ya məlumat sızdırır.

Səhlənkar daxili əməkdaşlar (Negligent insiders). Zərərli niyyət daşımadan, lakin diqqətsizlik, qaydalara əməl etməmək və ya təhlükəsizlik biliklərinin kifayət qədər olmaması səbəbilə insayder riskinin yaranmasına səbəb olan təşkilat əməkdaşlarıdır.

## III. İNSAYDER TƏHDİDLƏRİNİN NÖVLƏRİ

İnsayder təhdidləri aşağıdakı növlərə bölünür:

Qəsdən edilməyən insayder təhdidləri (unintentional insider threats). Bu təhdidlər diqqətsizlik, məlumatsızlıq və ya təsadüfi səhvlər nəticəsində yaranır. İstifadəçilər təhlükəsizlik qaydalarını bilsələr də, rahatlıq və ya səhlənkarlıq səbəbindən həmin qaydaları poza bilərlər. Məsələn, həssas sənədin səhv şəxsə göndərilməsi, zərərli linkə klik edilməsi və ya yoluxmuş faylın açılması bu kateqoriyaya aiddir.

Qəsdən edilən insayder təhdidləri (intentional insider threats). Şəxsi mənfəət əldə etmək və ya şəxsi narazılıq səbəbilə təşkilata zərər verən əməliyyatlarla əlaqəli insayder təhdididir. Buna görə qəsdən edilən insayderlərə "zərərli insayder" də deyilir. Bir çox insayder gözləntilərinin qarşılanmaması, məsələn, tanınmama, vəzifədə yüksəlməmə, maliyyə mükafatlarının verilməməsi və ya işdən çıxarılma kimi səbəblərə görə "qisas almağa" çalışa bilər. Belə zərərli insayderlər həssas məlumatları sızdıra, avadanlıqlara sabotaj edə, iş yoldaşlarını və ya əməkdaşları narahat edə bilər və s. Həmçinin onlar əqli mülkiyyəti və ya şəxsi məlumatları, müştəri məlumatlarını və digər məxfi informasiyanı oğurlaya bilərlər.

Razılaşdırılmış (əlbir) təhdidlər (collusive threats). Bir neçə insanın (məsələn, insayder və kənar hücumçunun) birlikdə əməkdaşlıq edərək təşkilata zərər vurması.

Üçüncü tərəf şəxslərin yaratdığı təhdidlər (third-party threats). Təşkilatın birbaşa işçisi olmayan, lakin sistemlərə və ya məlumatlara girişi olan xaricrdən (podratçılar, təchizatçılar, xidmət təminatçıları və s.) gələn təhdidlər.

## IV. KÖLGƏ SÜNİ İNTELLEKTİ VƏ İNSAYDER TƏHDİDLƏRİ

Verizon təşkilatının insayder təhdidləri ilə bağlı 2023-cü ildə hazırladığı hesabatda qeyd olunur ki, əməkdaşlar təşkilatın təhlükəsizlik siyasətlərini pozaraq icazəsiz süni intellekt alətlərindən istifadə edərək məlumat sızmaları yaradırlar. Hesabatda bu sızma növünün məlumat sızmalarının 55%-dən çoxunu təşkil etdiyi qeyd olunur [8]. Təşkilatın 2026-cı ildə nəşr etdirdiyi hesabatında isə insayder təhdidlərinin artmasına səbəb kölgə süni intellekti (Shadow AI) olduğu qeyd olunur [9]. Yəni əməkdaşlar şəxsi OpenAI hesabı ilə korporativ məlumatları, məxfi sənədləri ChatGPT tipli süni intellekt sistemlərinə daxil edirlər. Bu isə məlumatın üçüncü tərəf şəxslərə sızması risklərini artırır. Bu hesabatda kölgə süni intellekti təşkilatın məlumat sızmaları ilə bağlı 2025-ci ildə topladıqları məlumatlar içində üçüncü ən çox yayılmış zərərsiz insayder əməliyyatı kimi qeydiyyata alınmışdır.

## V. İNSAYDER TƏHDİDLƏRİNİN YARATDIĞI HÜCUMLAR

Karnegi Mellon Universitetinin CERT adlı Milli İnsayder təhdidləri mərkəzinin "Common Sense Guide to Mitigating Insider Threats" insayder təhdidlərinin yaratdığı hücumlara aşağıdakılar aid edilir [13].

Giriş imtiyazlarından sui-istifadə (Misuse of Access Privileges). İnsayderlərin onlara verilmiş qanuni giriş icazələrindən şəxsi məqsədlər və ya zərərli fəaliyyətlər üçün istifadə etməsidir.

Girişin idarə edilməsi mexanizminin sıradan çıxması (Access-control failure). Bu tip hücumda texniki problemlərin olması hücumun uğur qazanmasının əsas səbəbidir. Bu hücum növü giriş sistemi düzgün konfiqurasiya olunmadıqda reallaşa bilir və icazəsiz insayder təşkilatın şəbəkəsinə giriş əldə edir.

İntellektual mülkiyyətin oğurlanması (Intellectual property theft). Bu insident insayderlərin təşkilatın kommersiya sirlərini, məxfi məlumatlarını, mənbə kodunu və ya strateji planlarını oğurlamaq üçün onlara verilmiş giriş icazələrindən sui-istifadə etməsi zamanı baş verir.

İT sabotajı (IT sabotage). Bu insident insayderlərin biznes əməliyyatlarını pozmaq və ya maliyyə zərəri vurmaq məqsədilə təşkilatın sisteminə, məlumatlarına və ya şəbəkələrinə qəsdən zərər verməsi zamanı baş verir.

Müdafiə mexanizmlərindən yayınma (Bypassing Defenses). İnsayderlərin daxili giriş imkanlarından istifadə edərək təhlükəsizlik nəzarətini keçməsi və ya deaktiv etməsidir.

Casusluq (Espionage). İnsayderlərin və ya xarici aktorlarla əməkdaşlıq edən şəxslərin təşkilatın məxfi məlumatlarını gizli şəkildə toplaması və ya ötürməsi ilə xarakterizə olunan hücum növüdür. Bu hücumlar adətən kommersiya sirrlərinin, strateji planların və ya dövlət əhəmiyyətli məlumatların oğurlanmasına yönəlir.

## VI. MƏŞHUR İNSAYDER HÜCUMU FAKTLARI

İnsayder təhdidləri riskinin ciddi olduğunu qiymətləndirmək üçün aşağıda bir sıra məşhur insayder hücumu faktları göstərilmişdir [10]:

2023-cü ildə Tesla şirkətinin əməkdaşları tərəfindən minlərlə fərdi məlumat Alman Xəbər Agentliyinə (German News Outlet) sızdırılmışdır. Bu əməliyyatın nəticəsində Tesla şirkətinin iki keçmiş işçisi 75735 cari və keçmiş əməkdaşın adlarını, ünvanlarını, telefon nömrələrini və e-poçt ünvanlarını Alman qəzetinə sızdırmışdır.

2020-ci ildə General Electric şirkətinin işçiləri biznes üstünlüyü əldə etmək üçün kommersiya sirrlərini oğurlamışdır. Bu insident nəticəsində General Electric şirkətinin iki işçisi təşkilatın serverindən minlərlə ticarət sirri olan faylları yükləmişlər, onları buluda və ya öz e-poçt ünvanlarına göndərmişlər. Bu işçilər həmçinin təşkilatın həssas məlumatlarına giriş əldə etmək üçün sistem administratorunu razı salmışlar. Əldə etdikləri məlumatları özlərinin gələcək biznes fəaliyyətlərində istifadə etmişlər.

2019-cü ildə Capital One şirkətinin məlumatları Amazon Veb Xidmətlərinin proqram mühəndisi olan qadın tərəfindən sızdırılmışdır. Səhv konfiqurasiya edilmiş veb tətbiqinin şəbəkələrarası ekranından istifadə edən haker 100 milyondan çox müştərinin hesabına və kredit kartı tətbiqlərinə daxil olmuşdur.

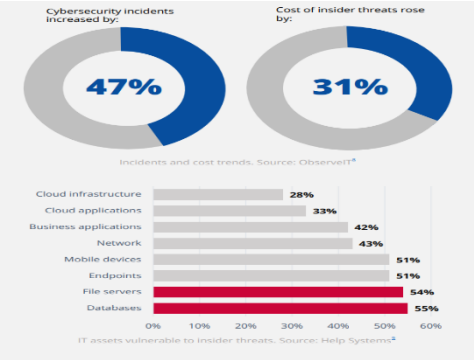
2018-ci ildə Tesla şirkəti ilə bağlı baş verən insidentdə hücumçu saxta istifadəçi adları altında Tesla İstehsalat Əməliyyat Sistemində birbaşa kod dəyişiklikləri edərək və çoxlu miqdarda yüksək həssas Tesla məlumatlarını naməlum üçüncü tərəflərə ötürmək üçün daxili giriş səlahiyyətindən istifadə etmişdir.

2015-ci ildə ABŞ-da fəaliyyət göstərən Anthem adlı böyük səhiyyə sığorta şirkətinin fərdi məlumatları hesaba müdaxilə insidenti nəticəsində sızdırılmışdır. Burada xakerlərin göndərdiyi fişinq e-poçtu nəticəsində 78.8 milyon fərdi identifikasiya məlumatları ələ keçirilmişdir. Bu insident nəticəsində tibbi identifikatorlar, sosial sığorta nömrələri, ünvanlar, məşğulluq tarixçələri, gəlir məlumatları oğurlanıb Dark web-də satılmışdır. Hakerlər işçilərinə zərərli proqram təminatına keçidlər olan fişinq e-poçtları göndəriblər. Kliklədikdən sonra zərərli proqram təminatı quraşdırılıb və hakerlər verilənlər bazasına arxa qapı açıblar. Onlar verilənlər bazasına uzaqdan daxil ola biliblər.

2013-cü ildə Target pərakəndə satış şirkətinin 41 milyondan çox müştəri ödəniş kartı hesabı oğurlanmışdır. Hakerlər üçüncü tərəf təchizatçıdan oğurlanmış giriş məlumatları vasitəsilə Target şirkətinin kompüter sistemlərinə daxil ola bilmişdir və zərərli proqram yükləyərək şirkətin müştərilərinin ad, soyad, telefon nömrəsi, e-poçt ünvanı, ödəniş kartının nömrəsi, kredit kartının təsdiqləmə kodunu, və s. kimi həssas məlumatlarını ələ keçirmişlər.

## VII. İNSAYDER TƏHDİDLƏRİNƏ QARŞI MÜBARİZƏ STRATEGİYALARI

ENISA təşkilatının hesabatında insayder təhdidlərinin sürətlə artığı qeyd edilmişdir (Şəkil 1) [11].




Şəkil 1. İnsayder təhdidlərinin artım dinamikası

Şəkildən göründüyü kimi təşkilatlarda baş verən təhlükəsizlik hadisələrinin sayı ciddi şəkildə yüksələrək əvvəlki illərlə müqayisədə 47% artmışdır. İşçilər, keçmiş əməkdaşlar və ya daxili giriş icazəsi olan şəxslərin yaratdığı problemlərin təşkilatlara vurduğu maddi zərər böyüyərək onların vurduğu maliyyə ziyanı 31% artmışdır. Bundan əlavə həmin sənəddə IT resursların insayder təhdidlərinə qarşı daha həssas olduğu sahələr müəyyən edilmişdir: Verilənlər bazaları (Databases) - 55%; Fayl serverləri (File servers) - 54%; Endpoints (kompüter və cihazlar) - 51%; mobil cihazlar - 51%; şəbəkə (Network) - 43%; biznes tətbiqləri-42%; bulud tətbiqləri - 33%; bulud infrastrukturu -28%. Buradan göründüyü kimi insayder təhdidləri ilə bağlı ən böyük risk məlumatların saxlandığı sistemlərdədir. Xüsusilə verilənlər bazaları, fayl serverləri, mobil cihazlar və istifadəçi kompüterləri də yüksək risk daşıyır.

İnsayder risklərinin sürətlə artımı, insayder təhdidlərindən müdafiə texnologiyalarının işlənməsini zəruri etmişdir. Ponemon İnstitutunun 2025-ci ildə dərc etdirdiyi hesabatda qeyd edilir ki, dünya təşkilatlarının 81 faizi artıq insayder təhdidlərinin idarə edilməsi proqramını tətbiq edir [6]. İnsayder təhdidlərinin idarə edilməsi proqramı həssas məlumatların icazəsiz şəkildə açıqlanmasını aşkarlamaq və qarşısını almaq üçün mərkəzləşdirilmiş idarəetmə altında təşkil olunmuş koordinasiyalı subyektlərin qrupudur [12]. Bu qrup təşkilatlara insayder təhdidlərini aşkarlamaqda, onlara cavab verməkdə, fəsadların aradan qaldırılmasında və insayder təhdidləri haqqında məlumatlılığı artırmaqda köməklik edə bilər.

## VIII. İNSAYDER TƏHDİDLƏRİNİN AŞKARLANMASI SAHƏSİNDƏ MÖVCUD SÜNİ İNTELLEKT YANAŞMALARI

İnsayder təhdidləri işçilərin identifikasiya məlumatlarını ələ keçirərək və həmin məlumatlardan istifadə etməklə hücumlar həyata keçirdiyindən onların aşkarlanması istifadəçi davranışının analizinə əsaslanmalıdır [14]. Burada hər bir istifadəçinin davranışı qeyd olunur və standart qaydalar çoxluğu ilə müqayisə olunur. İstifadəçinin davranışı normaldan kənara çıxdığı halda, onun davranışı zərərli kimi qiymətləndirilir.

İnsayderlərin istifadəçi davranışına əsasən aşkarlanması üçün çox sayda maşın təlimi və dərin təlim əsaslı yanaşmalar işlənmişdir. [15]-də istifadəçi davranışının analizi əsasında insayder hücumlarının aşkarlanması üçün LSTM-Autoencoder əsaslı yanaşma işlənmişdir. İstifadəçinin fəaliyyətinə uyğun olaraq istifadəçi davranışı normal və zərərli kimi kateqoriyalaşdırılmışdır. Zərərli davranışın aşkarlanması üçün

sistemə giriş-çıxış hadisələri, istifadəçi rolları kimi əlamətlər istifadə edilmişdir. İstifadəçinin fəaliyyətinə əsasən onun davranışı normal və zərərli siniflərdə kateqoriyalaşdırılmışdır. Eksperimentlərin aparılması üçün CMU CERT sintetik verilənlər bazası istifadə edilmişdir. [16]-da insayder hücumlarının aşkarlanması üçün çoxmiqyaslı qrup təlimi modeli işlənmişdir. Modeli üç komponent təşkil edir:

1) zamandan asılı (temporal) əlamətlərin kollaborativ çıxarılması modulu. Burada çəkili diqqət mexanizmi istifadə edilərək istifadəçilərin ayrı-ayrı davranışları əldə olunub və onların birləşdirilməsi həyata keçirilib.

2) qraf neyron şəbəkələrinin aqreqasiya mexanizmindən istifadə etməklə istifadəçilər arasında struktur asılılıq əldə olunub.

3) seyrək diqqət mexanizminin tətbiqi ilə istifadəçinin normal davranışından kənaraçıxma qiymətləndirilib. Yanaşmanın effektivliyinin qiymətləndirilməsi CERT r4.2 və CERT r5.2 verilənlər bazaları üzərində aparılıb.

[17]-də insayder hücumlarının aşkarlanması ilə bağlı işlənmiş maşın təlimi metodlarının test edilməsi üçün yararlı ola bilən SPEDIA adlı verilənlər bazası yaradılmışdır. Verilənlər bazası istifadəçilərin real davranışı, real işçi rollarının üzərində modelləşdirilmiş simulyasiya edilmiş istifadəçi davranışı, və CERT Insider Threat verilənlər bazasından götürülmüş sintetik verilənlər təşkil edir. Verilənlər bazası istifadəçi fəaliyyətinə aid hadisələrin loq yazılarını qeydə alıb. Buraya fayllar üzərində əməliyyatlar, əmrlərin icrası, xidmətlərdən istifadə və şəbəkə davranışı kimi əməliyyatlar daxildir.

[18]-də insayder təhdidlərinin nəzarətsiz aşkarlanması üçün LAAER adlı yeni yanaşma təklif edilib. Yanaşma izolyasiya ağacı tipli anomaliyaların aşkarlanması üsullarının insayder təhdidlərinin daha dəqiq aşkarlamasını təmin etmək üçün loq kontentdəki mətn tipli linqvistik məlumatdan istifadə edir. Müxtəlif ssenarilərdə zərərli fəaliyyətin tapılması üçün insayder təhdidlərinin diqqət, emosiya və davranış anomaliyaları adlı üç nümunəsindən istifadə edilib. Diqqət anomaliyalarının aşkarlanmasında əməliyyat obyektlərindəki (e-poçtlar, veb səhifələr) mətn kontentləri analiz edilərək təhdidlər aşkarlanır. Burada mətn məlumatları əməkdaşların diqqət yönəltdiyi sahələri əks etdirir. Diqqət gündəlik iş fəaliyyətindən ciddi şəkildə kənara çıxdıqda, əməkdaşın fəaliyyəti zərərli kimi qeydə alınır. İşçinin işə aid olmayan sahələrlə maraqlandığı mətn tipli məlumatları əhatə edir. Emosional anomaliyaların aşkarlanması, iki işçinin gündəlik ünsiyyət qurduğu mesajlar arasındakı bütün dialoqları təhlil edir və potensial psixoloji problemləri tapmağa çalışır. İşçilər arasında ifadə edilən anomal emosiyalardan ibarət linqvistik mətnləri əhatə edir. Məsələn, e-poçt göndərmə/qəbul etmə dialoqu zamanı istifadə edilən neqativ və ya depressiv sözlər. Davranış anomaliyasının aşkarlanması bloku təhdidləri aşkarlamaq üçün loq fayllarda qeydə alınan əməliyyatları analiz edir. İşçilərin gündəlik işlərində yol verdikləri qeyri-adi davranış tərzini əhatə edir. Təklif edilmiş yanaşma diqqət və emosiya anomaliyalarından əldə edilən məlumatlardan köməkçi əlamətlər kimi istifadə edir və bu əlamətləri loq əməliyyatlarından çıxarılan əlamətlərlə və statistik qiymətlərlə inteqrasiya edərək loq vektorları yaradır. LAAER yanaşması formalaşmış loq vektorlara anomaliyaların İzolyasiya Ağacı tipli aşkarlanması alqoritmlərini tətbiq edir və əməkdaşın zərərli əməliyyatlarını analiz edir. Daha sonra eyni şöbədə çalışan bütün əməkdaşları nəzərə almaqla əməkdaşın davranış anomaliyasını aşkarlayır. İki kontent arasındakı dəyişikliyin dərəcəsinin ölçülməsi üçün kosinuslar məsafəsi düsturundan istifadə edilib. Sözlərin vektor təsvirinin əldə edilməsi üçün təbii dilin emalı nəzariyyəsinin məşhur Glove, BERT, GPT-3 metodları istifadə edilir. İşdə Glove tətbiq olunaraq ölçüsü 50 olan söz vektoru əldə edilmişdir. İşdə LAAER modelinin prototipi yaradılmışdır və CERT və LANL verilənlər bazaları üzərində test edilmişdir.

[19]-da insayder təhdidlərinin aşkarlanmasında istifadəçilərlə obyektlər (məsələn, fayllar, proseslər, kompüterlər, veb-saytlar və çıxarıla bilən qurğular) arasındakı əlaqələri, eləcə də istifadəçi davranış nümunələrinin zamanla dəyişməsini nəzərə almaq üçün diqqət mexanizmi əsaslı temporal heterogen qraf neyron şəbəkəsi (Attention-based Temporal Heterogeneous İnsider Threat Detection, ATHITD) təklif edilmişdir. İlk olaraq, ATHITD istifadəçilər və obyektlər arasındakı dəyişən və mürəkkəb əlaqələri təsvir etmək məqsədilə müəyyən edilmiş zaman pəncərəsi əsasında müxtəlif loqlardan temporal heterogen qraf ardıcıllıqları qurur. Daha sonra, hər bir zaman pəncərəsində hədəf qovşaqlar üçün qısamüddətli temporal asılılıqları təsvir etmək məqsədilə temporal qonşular müəyyən edir. Temporal qonşular hədəf qovşaqlarla eyni olan və əvvəlki zaman pəncərələrində mövcud olmuş qovşaqlardır. Bundan sonra diqqət mexanizmindən istifadə edilərək hədəf qovşaqların məkan heterogenliyi və temporal qonşulardan hədəf qovşaqlara doğru qısamüddətli xüsusiyyət dəyişiklikləri öyrənilir. Əlavə olaraq transformer modelindəki self-attention mexanizmi vasitəsilə müxtəlif zaman pəncərələri boyunca istifadəçi qovşaqlarının uzunmüddətli xüsusiyyət dəyişiklikləri öyrənilir. ATHITD həmçinin zərərli fəaliyyətlərin baş verdiyi zaman pəncərələrinə fokuslana bilir ki, bu da təhlükəsizlik mütəxəssislərinə vurğulanmış zaman intervallarında potensial zərərli fəaliyyətləri analiz etməyə kömək edir. Yanaşmanın test prosesi CERT və LANL verilənlər bazaları üzərində aparılmışdır.

[20]-də insayder təhdidlərinin aşkarlanması üçün statistik və zamandan asılı ardıcıl məlumatların analizini həyata keçirən konvalyusiya diqqət mexanizmi və transformer kodlayıcısından ibarət hibrid model təşklif edilmişdir. Burada ilk olaraq, istifadəçilərin davranış loqları müxtəlif mənbələrdən toplanaraq sonrakı analiz üçün uyğun formata gətirilmişdir. Statistik məlumatları effektiv şəkildə öyrənmək üçün konvolyusiya əsaslı diqqət mexanizminə əsaslanan alt şəbəkə strukturu, ardıcıl məlumatların öyrənilməsi üçün isə transformer əsaslı alt şəbəkə strukturu qurulmuşdur. Təklif edilmiş yanaşmanın effektivliyi və dayanıqlığı CERT verilənlər bazası üzərində test edilmişdir. [4]-də istifadəçi davranışını əks etdirən təsvirlərin yaradılması üçün kontekstual əlamətlərin çıxarılmasını və anomal əlamətlərin gücləndirilməsini həyata keçirən diqqət modulu işlənmişdir. Əvvəlcədən öyrədilmiş ResNet və davranış, psixoloji və rollara əsaslanan əlamətlərdən ibarət çoxmənbəli əlamətlərin birləşdirilməsindən istifadə etməklə zərərli insayderlər identifikasiya olunmuşdur. Təklif olunmuş yanaşma CMU-CERT Insider Threats Dataset üzərində test edilmişdir.

[21]-də insayder təhdidləri anomaliyalarının aşkarlanması üçün avtoenkodlayıcı əsaslı nəzarətsiz dərin təlim yanaşması işlənmişdir. Təklif edilmiş yanaşma istifadəçilərin və cihazların

anomal davranışlarını analiz edir və təyin edilmiş sərhəd qiymətinə əsasən xəbərdarlıq göndərir. Təhlükəsizlik aşkarlanması üçün təyin edilmiş həddi qiyməti normal verilənlərin test edilməsi nəticəsində əldə olunan rekonstruksiya səhvləri əsasında hesablanır. Təklif olunan model istifadəçi davranışını yenidən qurarkən böyük rekonstruksiya xətası yaratmırsa, yəni xəta müəyyən edilmiş sərhədd qiymətini aşmırsa, istifadəçi normal kimi qəbul edilir, əks halda, istifadəçi insayder kimi qeydə alınır.

[22]-də insayder hücumlarının aşkarlanması üçün ansambl təlimi metodu təklif edilmişdir. Ansambl modelini Decision Tree, Naive Bayes, KNN, Logistic Regression və SVM alqoritmləri təşkil edir. Eksperimentlər CERT verilənlər bazası üzərində aparılmışdır.

[23]-də insayder hücumlarının ansambl və özünüidarəetmə (self-supervised) təliminə əsaslanan supervizorlu aşkarlanması metodu təklif edilmişdir. Aşkarlanmanın effektivliyinin artırılması üçün yanaşmaya subyektləri TF-IDF əsasında vektor şəklində təsvir edən metod daxil edilmişdir. Eksperimentlər CERT4.2 və CERT6.2 verilənlər bazaları üzərində edilmişdir.

[24]-də insayder hücumlarının aşkarlanması üçün Word2vec, GLoVe və LSTM modellərinin birləşdirilməsi əsasında yaradılmış hibrid model təklif edilmişdir. Təklif edilmiş yanaşmanın XGBoost, AdaBoost, RF (Random Forest), KNN (K-Nearest Neighbor) və Logistic Regression modelləri ilə müqayisəli analizi aparılmışdır. Eksperimentlərin aparılması üçün "Enron e-mail dataset" istifadə edilmişdir.

[25]-də davranış analizinə əsaslanan insayder təhdidlərinin aşkarlanması məsələsi araşdırılmışdır. Tədqiqat çərçivəsində verilənlərin toplanması, ilkin emalı, əlamətlərin seçilməsi və modelin qurulması mərhələlərindən ibarət kompleks aşkarlama strategiyası təklif edilmişdir. İnsayder təhdidlərinin müəyyən edilməsi üçün əsas model kimi diqqət mexanizmi ilə gücləndirilmiş LSTM arxitekturasından istifadə olunmuşdur.

[26]-də insayder təhdidlərinin aşkarlanması üçün təbii dil emalının qabaqcıl metodlarını, statistik əlamət analizini və psixometrik profilləşdirməni birləşdirən yeni metodologiya təklif olunub. Yanaşmada kompleks davranış profillərini formalaşdırmaq üçün istifadəçilərin kommunikasiya və veb fəaliyyətlərindəki mətn kontenti, zaman və tezlik əsaslı davranış göstəricilərinin statistik qiymətləri, həmçinin OCEAN göstəricilərinə əsaslanan psixometrik xallari birləşdirilmişdir. Şübhəli və normal davranışların müəyyən edilməsi üçün izolyasiya ağacı (isolation forest) modeli istifadə edilmişdir. Əlamətlərin vacibliyini müəyyən etmək üçün ölçünün azaldılması üsulu tətbiq edilmişdir.

[27]-də insayder təhdidlərinin istifadəçi davranışının analizi əsasında aşkarlanması üçün təkmilləşdirilmiş yanaşma təklif edilib. Təklif olunan yanaşma bir neçə mərhələdən ibarətdir: küyün aradan qaldırılması məqsədilə verilənlərin ilkin emalı, böyük ölçülü fəzadan əlamətlərin çıxarılması zamanı informasiya itkisinin minimallaşdırılması üçün izometrik əlamət xəritələndirməsinin (Isometric Feature Mapping) tətbiqi, yekun klassifikasiyanın dəqiqliyinin artırılması üçün kontent əsaslı əlamətlərin inteqrasiyası, optimal əlamət seçimi üçün effektiv axtarış və optimallaşdırma qabiliyyətinə malik Emperor Penguin alqoritminin tətbiqi, həmçinin müxtəlif növ əlamətlərin paralel emalı və sürətli hesablanması üçün çoxsəviyyəli qeyri-səlis (multi-fuzzy) klassifikatorun tətbiqi. Yanaşmanın effektivliyi CMU-CERT v4.2 verilənlər bazası üzərində test edilmişdir və səkkiz merika əsasında qiymətləndirilmişdir.

[28]-də insayder təhdidlərinin aşkarlanması üçün çoxsaylı aşkarlama metodlarını birləşdirən hibrid süni intellekt üsulu təklif olunub. Yanaşmada anomaliyaların nəzarətsiz, nəzarətli, ansambl və dərin təlim əsasında aşkarlanması arxitekturaları birləşdirilib. Verilənlərə əlamət mühəndisliyi tətbiq olunaraq zaman, davranış, şəbəkə və psikometrik indikatorlar əldə edilmişdir. İsolation Forest və Local Outlier Factor metodları tətbiq edilərək verilənlərdən anomaliya nişanları yaradılmış və hibrid modelin öyrədilməsi təmin edilmişdir. Yanaşmanın effektivliyi "CERT 4.1 insider threat dataset" üzərində test edilərək ROC-AUC, precision, recall, F1-score, və cost sensitive analysis metrikaları əsasında qiymətləndirilmişdir. SHAP metodunun diqqət (attention) mexanizmləri ilə birləşdirilmiş forması nəticələrin interpretasiyası üçün istifadə olunmuşdur və hər bir əlamətin vaciblik dərəcəsi göstərilmişdir.

[29]-da Əşyaların İnternetinin (Internet of Things, IoT) xarakteristikalarına və strukturuna əsasən insayder təhdidlərinin müxtəlif aspektləri analiz olunub. IoT-un qavrama, şəbəkə, tətbiqi laylarında toplanan verilənlərin insayder hücumlarının aşkarlanmasında səmərəliliyi qiymətləndirilib. Hər bir layın xarakteristikalarına əsasən onların uyğun verilənlər mənbələrinin klassifikasiyası aparılıb, hər bir kateqoriya üzrə tədqiqatın məqsədi və metodları tədqiq edilib.

[30]-da insayder təhdidlərinin aşkarlanması üçün maşın təlimi və dərin təlimə əsaslanan yanaşma təklif edilmişdir. Yanaşmada istifadəçilərin və ya kiber-şəxs (cyber persona) adlanan istifadəçilər qrupunun normal davranışı öyrənilir, normal davranışdan kənaraçıxmalar müəyyən olunur və onlar insayder təhdidləri kimi qeydiyyata alınır. Yanaşma bir neçə komponentdən ibarətdir. Burada istifadəçinin həyata keçirdiyi əməliyyatlarını maşının həyata keçirdiyi əməliyyatlardan fərqləndirmək üçün Machine/Human Segregation modulu nəzərdə tutulmuşdur. Bu modul Feed Forward Neural Networks (FFNN), Convolutional Neural Networks (CNN), LSTM, decision trees, random forest kimi bir neçə binar klassifikatordan ibarətdir. Növbəti blokda istifadəçi ilə bağlı verilənlər emal olunaraq onların şəxsiyyətinin tipi identifikasiya olunur. Bu məqsədlə verilənlərə yuxarıda adları çəkilən maşın təlimi alqoritmləri tətbiq olunur. Sonuncu insayder təhdidlərinin aşkarlanması modulunda istifadəçinin davranış kənaraçıxmaları analiz olunaraq insayder təhdidləri işarələnir. Bu modulda izolyasiya meşələri (isolation forests), dayanıqlı kovariasiya (robust covariance), birsinifli dayaq vektorları metodu (one-class Support Vector Machine, SVM), avtokodlayıcı (autoencoder) tipli kənaraçıxmaların aşkarlanması üsulları tətbiq olunur. Eksperimentlərin aparılması üçün müəlliflər tərəfindən yaradılmış verilənlər bazaları istifadə edilmişdir.

[31]-də insayder hücumlarının aşkarlanması üçün təklif edilmiş yanaşmada zamandan asılı əlamətlərin çıxarılması üçün LSTM, aşkarlanmanı həyata keçirmək üçün isə CNN istifadə edilib.

İnsayder təhdidlərinin aşkarlanması və qarşısının alınması sahəsində mövcud yanaşmaların təhlilinin nəticələri Cədvəl 1-ə daxil edilmişdir.

CƏDVƏL I. MÖVCUD YANAŞMALARIN TƏHLİLİNİN NƏTİCƏLƏRİ

| Ədəbiyyat | Metod | Verilənlər bazası | Məqsəd |
|---|---|---|---|
| [16] | MSGL - multi-scale group learning model | CERT r4.2, CERT r5.2 | Anomaliyaların aşkarlanması |
| [17] | Maşın təlimi modellərinin yaradılmış verilənlər bazaları üzərində test edilməsi | SPEDIA, CERT Insider Threat | İnsayder təhdidləri ilə bağlı verilənlər bazasının yaradılması |
| [18] | Cosine similarity, Glove, BERT, GPT-3, isolation forest | CERT, LANL | Anomaliyaların aşkarlanması |
| [19] | Diqqət mexanizmli transformer modeli | CERT, LANL | İnsayder təhdidlərinin modelləşdirilməsi |
| [21] | Avtokodlayıcı | Müəlliflərin topladığı verilənlər bazası | Anomaliyaların aşkarlanması |
| [23] | Özünənəzarətli təlim, TF-IDF | CERT4.2, CERT6.2 | Anomaliyaların aşkarlanması |
| [24] | Word2vec, GLoVe və LSTM | Enron e-mail dataset | Davranış analizi |
| [25] | Diqqət mexanizmi ilə gücləndirilmiş LSTM | Müəlliflərin topladığı verilənlər bazası | Davranış analizinə əsaslanan insayder təhdidlərinin aşkarlanması |
| [26] | NLP, Isolation forest | CERT r4.2 | Davranış profilindən kənaraçıxmaların müəyyən edilməsi |
| [27] | Isometric Feature Mapping, Emperor Penguin Algorithm | CMU-CERT v4.2 | Anomaliyaların aşkarlanması |
| [30] | Isolation forests, robust covariance, one-class SVM, autoencoder | Müəlliflərin topladığı verilənlər bazası | İnsayder təhdidi rolunda çıxış edən kiber-şəxsin müəyyən edilməsi |

Cədvəldən göründüyü kimi, insayder təhdidlərinin aşkarlanması sahəsində aparılmış tədqiqatlarda əsasən CERT, LANL və Enron kimi verilənlər bazalarından istifadə edilmişdir. Mövcud yanaşmalarda anomaliyaların aşkarlanması və davranış analizinə əsaslanan müxtəlif maşın təlimi və dərin təlim metodları tətbiq edilmişdir. Təhlil göstərir ki, son illərdə transformer, avtokodlayıcı, NLP və özünənəzarətli təlim kimi müasir yanaşmalar insayder təhdidlərinin daha effektiv aşkarlanması məqsədilə geniş istifadə olunur.

## IX. İSTİFADƏÇİ DAVRANIŞI ANALİZİ VƏ İNSAYDER TƏHDİDLƏRİNİN AŞKARLANMASI ÜÇÜN MÖVCUD ALƏTLƏR

İstifadəçi Davranışının Analizi (User Behavior Analytics, UBA) termini ilk dəfə 2015-ci ildə Gartner təşkilatı tərəfindən işlədilmişdir [32]. İstifadəçi davranışı analizi sistemləri üç yanaşma əsasında hazırlana bilər:

- Məlumat mərkəzli yanaşma (data-centric approach);
- İnsan mərkəzli yanaşma (user-centric approach);
- Sistem mərkəzli yanaşma (system-centric approach).

UBA daxili təhdidlərin, məqsədli hücumların və maliyyə fırıldaqçılığının aşkarlanması üçün yaradılmış kiber təhlükəsizlik alətidir. Bu alət potensial təhdidlərə işarə edən anomaliyaların aşkarlanması üçün xüsusiləşdirilmiş alqoritmlər və statistik analiz tətbiq etməklə istifadəçilərin davranış obrazlarını analiz edir. İnsayder təhdidlərinin aşkarlanması məsələsinin əsası hesab olunur.

İnsayder təhdidlərinin aşkarlanması üçün kiber təhlükəsizlik sahəsində fəaliyyət göstərən bir sıra aparıcı təşkilatlar tərəfindən alətlər hazırlanmışdır. Bu alətlər istifadəçi və obyek davranışının analizinə əsaslanırlar:

- ArcSight Intelligence. OpenText şirkətinə məxsus istifadəçi və obyekt davranışının analitikası (User and Entity Behavior Analytics, UEBA) həllidir. İstifadəçilərin və sistemlərin rəqəmsal izlərini izləmək üçün supervizorsuz maşın təlimi metodlarından istifadə edir və gündəlik milyardlarla təhlükəsizlik hadisəsini prioritetləşdirilmiş təhlükə siqnallarına çevirərək insayder təhdidlərinin və komprometasiya olunmuş hesabların aşkarlanmasını təmin edir [33].
- IBM QRadar User Entity Behavior Analytics. UEBA platformasıdır, insayder təhdidlərinin, eləcə də komprometasiya olunmuş hesabların aşkarlanması üçün istifadə olunur. Sistem normal davranış nümunələrini modelləşdirmək və ənənəvi qayda əsaslı sistemlərin müəyyən edə bilmədiyi anomal davranışları aşkarlamaq üçün maşın təlimi metodlarından istifadə edir [34].
- Kaspersky Endpoint Security. İstifadəçi və proses davranışlarını izləyərək anomal davranışları aşkarlayır.
- Kaspersky Security Center. Son istifadəçilərin informasiya sistemində baş verən təhlükəsizlik hadisələri haqqında məlumatları toplayır, istifadəçi davranışı ilə bağlı loq faylları analiz edir və şübhəli davranışları müəyyən edərək monitorinq həyata keçirir.
- Microsoft Sentinel. Bulud əsaslı SIEM və SOAR platforması olub, təhlükəsizlik hadisələrinin toplanması, analiz edilməsi və avtomatik cavablandırılması funksiyalarını həyata keçirir. Sistem maşın təlimi və davranış analizi metodlarından istifadə edərək anomal fəaliyyətləri, komprometasiya olunmuş hesabları və potensial insayder təhdidlərini aşkarlaya bilir [35].
- Splunk UBA. Anomal istifadəçi davranışlarını və mümkün təhdidləri aşkarlamaq üçün maşın təlimi və statistik üsullardan istifadə edir [36].
- Securonix UEBA. İstifadəçi davranışlarının müəyyən edilməsi və həqiqi əməliyyatları potensial təhdidlərdən ayırmaq üçün qabaqcıl maşın təlimi istifadə edir. Alət daxili təhdidləri və xarici müdaxilələri effektiv müəyyən etmək üçün real-vaxt məlumatlarını keçmiş davranış təhlili ilə birləşdirir və minimal küylə mürəkkəb təhdid aşkarlamasını təmin edir [37] .
- Exabeam UEBA. İstifadəçi, obyekt və süni intellekt agentləri ilə əlaqəli təhdidlərin aşkarlanması, araşdırılması və qarşısının alınması məqsədilə təhlükəsizlik əməliyyatları komandasına kömək edən qabaqcıl UEBA həllidir. Sistem Agent Davranış Analitikası (Agent Behavior Analytics, ABA) vasitəsilə süni intellekt və avtomatlaşdırma texnologiyalarından istifadə edərək təhlükəsizlik hadisələrini analiz edir [38].
- Varonis data-centric UEBA. Məlumat mərkəzli yanaşma tətbiq edir. Son nöqtələri izləmək əvəzinə,

istifadəçilərin bulud və yerli sistemlərdə məlumatlarla necə qarşılıqlı əlaqədə olduğunu izləyir [39].

- Forcepoint UEBA. Real obyekt risklərinin müəyyən edilməsi üçün süni intellekt əsaslı davranış analitikasıdır [40].

## NƏTİCƏ

Aparılmış təhlil göstərir ki, insayder təhdidlərinin aşkarlanması sahəsində mövcud yanaşmalar əsasən istifadəçi davranışlarının statistik xarakteristikalarının analizinə əsaslanır və əksər hallarda loq yazılarında mövcud olan mətn tipli məzmun məlumatlarını, istifadəçilərlə müxtəlif obyektlər arasındakı qarşılıqlı əlaqələri, eləcə də davranış nümunələrinin zamanla dinamik dəyişməsini kifayət qədər nəzərə almır. Bundan əlavə, kiber təhlükəsizlik sahəsində insayder təhdidlərinin tədqiqi üçün çoxsaylı istifadəçi davranışlarını əhatə edən açıq verilənlər bazalarının məhdud olması mövcud metodların effektivliyinin artırılmasını çətinləşdirir. Süni intellekt və maşın öyrənməsi sahəsində əldə olunan sürətli inkişafa baxmayaraq, bu texnologiyaların kiber təhlükəsizlik sistemlərində real tətbiqi hələ də məhdud səviyyədə qalır. Bununla yanaşı, istifadəçi davranışlarının daha dərin analizi, insan və avtomatlaşdırılmış bot fəaliyyətlərinin fərqləndirilməsi, DDoS hücumları və məzmunun avtomatik toplanması (content scraping) kimi zərərli fəaliyyətlərin aşkarlanması üçün intellektual metodların tətbiqi böyük əhəmiyyət kəsb edir. Bu baxımdan, gələcək tədqiqatlarda çoxmənbəli davranış məlumatlarını, mətn tipli loq analizini və zamanla dəyişən istifadəçi davranış modellərini nəzərə alan süni intellekt əsaslı yanaşmaların işlənməsi insayder təhdidlərinin daha dəqiq aşkarlanmasına imkan yarada bilər.

## MİNNƏTDARLIQ

Bu iş Azərbaycan Texniki Universitetinin maliyyələşdirdiyi AzTU-DQL-2025-M01/62562515 nömrəli qrant layihəsi çərçivəsində hazırlanmışdır.

## ƏDƏBİYYAT

[1] Data Breach Investigations Report, 2023, 89 p.

[2] D. Cappelli, A. Moore, R. Trzeciak, "The CERT Guide to Insider Threats: How to Prevent, Detect, and Respond to Information Technology Crimes (Theft, Sabotage, Fraud)," Cisco press, 2012, 430 p.

[3] Insider risk trend report 2026, Signpost Six, 30 p.

[4] S.M. Duan, J.T. Yuan, B. Wang, "Contextual feature representation for image-based insider threat classification," Computers & Security, 2024, Vol. 140, pp. 103779.

[5] Cost of Insider Threats Global Report, Ponemon Institute, 2022, 45 p.

[6] Insider Threat Statistics for 2025: Key Facts, Types of Incidents, and Costs, August 06, 2025. https://www.syteca.com/en/blog/insider-threat-statistics-facts-and-figures.

[7] Insider Threat Statistics: Costs, Trends & Key Data, 2026, https://app.stationx.net/articles/insider-threat-statistics

[8] Data Breach Investigations Report, Verizon, 2023, 89 p.

[9] Data Breach Investigations Report, Verizon, 2026, 121 p.

[10] T. Landman, "Famous Insider Threat Cases. September," Techstrong, 2019.

[11] "Insider threat ENISA Threat Landscape," ENISA, 2019, 18 p.

[12] NIST SP 800-53 Rev. 5, Security and Privacy Controls for Information Systems and Organizations, 2020, 492 p.

[13] Common Sense Guide to Mitigating Insider Threats, Seventh Edition, Software Engineering Institute, Carnegie Mellon University, 2022, 175 p.

[14] Cost of ınsıder threats global report, Ponemon Institute, 2022, 45 p.

[15] N. Rida, A. Mehreen, L. Rabia, I. Waseem, "Behavioral based ınsider threat detection using deep learning," IEEE Access, Vol. 9, 2021, pp. 143266-143274

[16] M. Pang, W. Ou, W. Meng, et al. "MSGL: A multi-scale group learning model for insider threat detection," Expert Systems with Applications, 2026, Vol. 323, pp.

[17] D.Á Muñiz, L.P. Miguel, A.M. Muñoz, et al. "Design and generation of a dataset for training insider threat prevention and detection models: The SPEDIA dataset," Computers & Security, Vol. 161, 2026, pp. 104743.

[18] K. Fei, J. Zhou, Y. Zhou, et al. "LaAeb: A comprehensive log-text analysis based approach for insider threat detection," Computers & Security, Volume 148, 2025, pp. 104126.

[19] Y. Qi, C. Yan, Z. Wang, et al. "ATHITD: Attention-based temporal heterogeneous graph neural network for insider threat detection," Computers & Security, 2025, Vol. 157, pp. 104587.

[20] H. Xiao, Y. Zhu, B. Zhang, et al. "Unveiling shadows: A comprehensive framework for insider threat detection based on statistical and sequential analysis," Computers & Security, 2024, Vol. 138, pp. 103665.

[21] K. Saminathan, S. Mulka, S. Damodharan, et al. "An Artificial Neural Network Autoencoder for Insider Cyber Security Threat Detection," Future Internet, Vol. 15, No. 12, 2023, pp. 1-29.

[22] K. Sivanagireddy, S. Jagadeesh, K.S. Kumar, et al. "Detecting and Mitigating Insider Threat Attacks in Cloud using Machine Learning," Proc. of the IEEE 6th International Conference on Cybernetics, Cognition and Machine Learning Applications, 19-20 October 2024, Hamburg, Germany.

[23] C. Zhang, S. Wang, D. Zhan, et al. "Detecting Insider Threat from Behavioral Logs Based on Ensemble and Self-Supervised Learning," Security and communication networks, 2021, Vol. 2021, pp. 1-11.

[24] M.A. Haq, M.A. Khan, M. Alshehri, "Insider Threat Detection Based on NLP Word Embedding and Machine Learning," Intelligent Automation & Soft Computing, 2022, vol.33, no.1, pp. 619-635.

[25] M. Rajesh, V. Mohan, "Intelligent insider threat detection based on behavioral analysis," AIP Conf. Proc., 2026, Vol. 3345, Issue 1, pp. 020311.

[26] F.R. Alzaabi, A. Mehmood, "Classifying Malicious Insider Threats Based On User Activity and Behavioral Profile Using Machine Learning," Proc. of the International Conference on Engineering and Emerging Technologies, 27-28 December, 2024, Dubai, United Arab Emirates.

[27] M. Singh, B. Mehtre, S. Sangeetha, "User behavior based Insider Threat Detection using a Multi Fuzzy Classifier," Multimedia Tools and Applications, 2022, Vol. 81, pp. 22953-22983.

[28] J. Roberto, An Artificial Neural Network Autoencoder for Insider Cyber Security Threat Detection, Master of Science in Data Analytics, 2025, 64 p.

[29] A. Kim, J. Oh, J. Ryu, K. Lee, "A Review of Insider Threat Detection Approaches With IoT Perspective," IEEE Access, Vol. 8, pp. 78847-78867.

[30] B. Racherache, P. Shirani, A. Soeanu, "CPID: Insider threat detection using profiling and cyber-persona identification," Computers & Security, Vol. 132, 2023, pp. 103350.

[31] F. Yuan, Y. Cao, Y. Shang, et al. "Insider threat detection with deep neural network," Proc. of the 18th International Conference on Computational Science, Wuxi, China, June 11–13, 2018, Part I, pp. 43– 54.

[32] What is User and Entity Behavior Analytics (UEBA)?, https://www.ibm.com/think/topics/ueba

[33] ArcSight Intelligence Behavioral Analytics, Opentext, Cybersecurity, 2025, 2 p.

[34] IBM QRadar User Entity Behavior Analytics app 5.0.1, User Guide, 2025, 300 p.

[35] Microsoft Sentinel, Architecture options for MSSPs System4u SecuRadar, 2025, 19 p.

[36] Splunk User Behavior Analytics, Splunk, 2025, 2 p. https://www.splunk.com/en_us/resources/splunk-user-behavior-analytics-product-brief.html

[37] Securonix User and Entity Behavior Analytics (UEBA), Securonix, 2026. https://www.securonix.com/products/ueba/

[38] Exabeam User and Entity Behavior Analytics, UEBA, https://www.exabeam.com/capabilities/ueba/.

[39] Data-centric UEBA, Varonis, 2026, https://www.varonis.com/platform/data-centric-ueba

[40] Forcepoint User and Entity Behavior Analytics, UEBA, 2026, https://www.guardsense.com/ueba.asp

## Review of Artificial Intelligence Methods for Insider Threat Detection

**Fargana Abdullayeva**
Institute of Information Technology, Baku, Azerbaijan
*a_farqana@mail.ru*

**Abstract**- In the field of cybersecurity, some of the most dangerous attacks originate not from external attackers, but from insider threats caused by trusted users within the system.

Since such users possess authorized access to system resources, these attacks can easily bypass security mechanisms and are difficult to detect. This paper analyzes existing approaches for insider threat detection and prevention. In particular, user behavior analysis, feature selection, and the application of machine learning and deep learning methods are extensively investigated, while the types of insider attacks and their countermeasure mechanisms are analyzed. The study also identifies existing challenges, open research issues, and future development perspectives, and provides relevant recommendations.

**Keywords**- Insider threat detection; log file analysis; user and entity behavior analytics; deep learning; machine learning.