AZƏRBAYCAN RESPUBLİKASI

Əlyazması hüququnda

ZƏRƏRLİ PROQRAM TƏMİNATININ AŞKARLANMASI ÜÇÜN İNTELLEKTUAL ANALİZ METOD VƏ ALQORİTMLƏRİNİN İŞLƏNİLMƏSİ

İxtisas: 3339.01 – İnformasiyanın mühafizəsi üsulları və sistemləri, informasiya təhlükəsizliyi

Elm sahəsi: Texnika elmləri

İddiaçı: **Elşən Oruc oğlu Bağırov**

Fəlsəfə doktoru elmi dərəcəsi
almaq üçün təqdim edilmiş dissertasiyanın

AVTOREFERATI

Bakı – 2026

Dissertasiya işi Azərbaycan Respublikası Elm və Təhsil Nazirliyinin İnformasiya Texnologiyaları İnstitutunda yerinə yetirilmişdir.

Elmi rəhbər: Texnika elmləri doktoru, dosent
Yadigar Nəsib oğlu İmamverdiyev

Rəsmi opponentlər: Texnika elmləri doktoru, professor
Cavanşir Firudin oğlu Məmmədov

Texnika elmləri doktoru, professor
Ramin Rza oğlu Rzayev

Texnika üzrə fəlsəfə doktoru
Lalə Hekayət qızı Kərimova

Azərbaycan Respublikasının Prezidenti yanında Ali Attestasiya Komissiyasının Azərbaycan Respublikası Elm və Təhsil Nazirliyinin İnformasiya Texnologiyaları İnstitutunun nəzdində fəaliyyət göstərən FD 1.35 Dissertasiya Şurası

Dissertasiya şurasının sədri:
AMEA-nın həqiqi üzvü, texnika elmləri doktoru, professor
Rasim Məhəmməd oğlu Əliquliyev

Dissertasiya şurasının elmi katibi:
texnika elmləri doktoru, dosent
Zərifə Qasım qızı Cəbrayılova

Elmi seminarın sədri:
texnika elmləri doktoru
Mütəllim Mirzəəhməd oğlu Mütəllimov

İŞİN ÜMUMİ XARAKTERİSTİKASI

Mövzunun aktuallığı və işlənmə dərəcəsi. İnformasiya texnologiyalarının inkişafı və rəqəmsal sistemlərin geniş tətbiqi informasiya təhlükəsizliyinə təhdidlərin miqyasını və mürəkkəbliyini artırmışdır. Zərərli proqramların yeni növləri və təkmilləşən yayınma üsulları ənənəvi mühafizə vasitələri qarşısında yeni tələblər qoyur. Buna görə onların vaxtında və etibarlı aşkarlanması üçün intellektual analiz metod və alqoritmlərinin işlənməsi mühüm elmi və praktiki əhəmiyyət daşıyır.

Ölkəmizdə kibermüdafiə çərçivəsində mühüm addımlar atılmağa başlanmışdır. Belə addımlardan biri olan Azərbaycan Respublikası Prezidentinin 2023-cü il 28 avqust tarixli Sərəncamı ilə təsdiq edilmiş "Azərbaycan Respublikasının informasiya təhlükəsizliyi və kibertəhlükəsizliyə dair 2023 – 2027-ci illər üçün Strategiyası" müasir İnformasiya Kommunikasiya Texnologiyalarından (İKT) təhlükəsiz istifadənin təmin olunmasına, kritik infrastrukturların qorunmasına və fərdi məlumatların mühafizəsinə yönəlmişdir[1]. Həmin strategiyada nəzərdə tutulmuş hədəf göstəricilərdən biri də məhz informasiya mühafizəsinin texniki, kriptoqrafik, proqram və digər vasitələr üzrə təminatının təkmilləşdirilməsidir. Bu hədəfə nail olmaq üçün qarşıya qoyulmuş məqsədlərdən biri informasiya təhlükəsizliyi və kibertəhlükəsizlik sahəsində elmi tədqiqat və layihə-təcrübə işlərinin aparılmasına dövlət səviyyəsində davamlı təşkilati dəstək verilməsinin təmin olunmasıdır.

Zərərli proqramlar kompüter sistemlərində və rəqəmsal mühitdə icazəsiz və əsasən bəd niyyətli əməllərin reallaşdırılması məqsədilə yaradılan və istismar edilən bütün növ proqramları özündə əks etdirir[2].

Zərərli proqram təminatının aşkarlanması istiqamətində tədqiqatların aparılmasını zəruri edən digər amillər aşağıda sadalanmışdır:

---

[1] "Azərbaycan Respublikasının informasiya təhlükəsizliyi və kibertəhlükəsizliyə dair 2023–2027-ci illər üçün Strategiyası"nın təsdiq edilməsi haqqında Azərbaycan Respublikası Prezidentinin Sərəncamı // 28 avqust 2023-cü il, № 4060. – Bakı, – 2023.

[2] Əliquliyev, R.M. İnformasiya təhlükəsizliyi insidentləri / R.M. Əliquliyev, Y.N. İmamverdiyev. – Bakı: İnformasiya Texnologiyaları, – 2012. – 219 s.

1. Statistik rəqəmlərə əsasən zərərli proqram nümunələrinin, o cümlədən, yeni növ zərərli proqram təminatının sayında eksponensial artımın müşahidə edilməsi: AV-TEST təşkilatı tərəfindən gün ərzində 450 mindən çox zərərli proqram qeydə alınır[3].

2. Zərərli proqram təminatının destruktiv nəticələrinin yüksək olması: Zərərli proqramlar kompüterdə istifadəçinin bəzi əməliyyatlarına mane olmaq, işini ləngitməklə yanaşı, böyük miqdarda maliyə itkisi, hətta kritik infrastrukturlarda insan tələfatı kimi təzahürlərlə müşahidə edilə bilər.

3. Zərərli proqram təminatının mövcud antiviruslar tərəfindən aşkarlanmasında problemlərin mövcudluğu: Hazırda bir çox kommersiya məqsədli antivirus sistemləri siqnaturaya (proqramı identifikasiya edən unikal imza) etibar edir və yalnız öz bazasında mövcud olan zərərli proqramları "tanıya" bilir. Yeni meydana çıxan zərərli proqram təminatını aşkarlaması üçün kompüter sistemi ən azı bir dəfə zərər görür.

4. Rəqəmsal sistemlərin təhlükəsizliyində proqram mühəndisliyinin rolu: Rəqəmsal sistemlərin mürəkkəbləşməsi proqram təminatının təhlükəsizliyinə qoyulan tələbləri artırır. Təhlükəsizlik prinsipləri layihələndirmə mərhələsindən istismara qədər nəzərə alınsa da, zərərli proqramların sistemə nüfuz etmə riskini tam aradan qaldırmaq mümkün olmur. Bu səbəbdən dayanıqlı arxitekturalarla yanaşı, zərərli proqramları vaxtında aşkarlayan intellektual analiz metodlarının və alqoritmlərinin işlənib hazırlanması aktuallıq kəsb edir.

5. İntellektual analiz metodlarının zərərli proqram təminatının aşkarlanmasına tətbiqi imkanlarının mövcud olması: Texnologiyanın sürətli inkişafı zərərli proqram təminatının aşkarlanması istiqamətində tədqiqatların gücləndirilməsinə təsirsiz ötüşməmişdir. Maşın öyrənmə, dərin öyrənmə kimi intellektual analiz metodlarının müxtəlif məsələlərə tətbiqi daha effektiv yanaşmaların inkişafına yol açmışdır.

Tədqiqatın obyekti və predmeti. Tədqiqatın obyekti zərərli proqram təminatının aşkarlanması, tədqiqatın predmeti isə zərərli proqram təminatının aşkarlanması üçün intellektual analiz metod və alqoritmləridir.

---

[3] AV-TEST. AV-ATLAS: Malware statistics: [Electronic resource]. – URL: https://portal.av-atlas.org/malware?lang=en.

Tədqiqatın məqsəd və vəzifələri. Tədqiqat işinin məqsədi zərərli proqram təminatının yüksək dəqiqliklə aşkarlanması üçün intellektual analiz metod və alqoritmlərinin işlənməsidir. Tədqiqat zamanı aşağıdakı məsələlər qarşıya qoyulmuşdur:

- Zərərli proqram təminatının aşkarlanması üzrə elmi tədqiqat və yanaşmaların sistemli təhlili, onların işləmə prinsiplərinin, üstünlük və məhdudiyyətlərinin müqayisəli qiymətləndirilməsi, elmi və texnoloji boşluqların müəyyənləşdirilməsi;
- Zərərli proqram təminatının aşkarlanması prosesinin əsas mərhələlərini, komponentlərini və onlar arasındakı qarşılıqlı əlaqələri əks etdirən konseptual modelin işlənilməsi;
- Zərərli proqram təminatının aşkarlanmasına sərf olunan hesablama vaxtının minimallaşdırılması məqsədilə tərsinə mühəndislik yanaşması əsasında statik və dinamik analiz nəticəsində əldə edilən əlamətlər içindən informativ əlamətlərin seçilməsi metodunun işlənilməsi;
- Zərərli proqram təminatının klassifikasiya və klasterləşdirmə məsələlərinin həlli üçün maşın öyrənməsi alqoritmlərinə əsaslanan aşkarlama metodlarının işlənilməsi;
- Zərərli proqram təminatının yüksək dəqiqliklə aşkarlanması məqsədilə dərin öyrənmə arxitekturalarına əsaslanan metodun işlənilməsi;
- Generativ Rəqib Şəbəkələrdən istifadə etməklə təlim verilənlərinin zənginləşdirilməsi və zərərli proqram təminatının aşkarlanması modelinin dəqiqlik və ümumiləşdirmə qabiliyyətinin artırılmasına yönəlmiş metodun işlənilməsi;
- Təklif olunan zərərli proqram aşkarlama metodlarının izahediləbilənliyinin təhlili, dayanıqlığının, etibarlılığının və dəyişən real mühitə uyğunlaşma qabiliyyətinin eksperimental qiymətləndirilməsi.

Tədqiqat metodları. Qarşıya qoyulmuş məsələlərin həlli üçün maşın öyrənmə üsullarına aid olan klassifikasiya və klasterləşdirmə metodları, avtoenkoderlər, Generativ Rəqib Şəbəkələr və dərin öyrənmə metodlarından istifadə edilmişdir.

Müdafiəyə çıxarılan əsas müddəalar:

- Zərərli proqram təminatının statik və dinamik analizi nəticəsində formalaşdırılmış əlamətlər çoxluğundan informativ əlamətlərin seçilməsi üçün əlamətlərin diskret və kəsilməzliyi, modelin xəttiliyi xüsusiyyətlərini nəzərə alan hibrid filtrasiya metodu;
- Zərərli proqram təminatının aşkarlanması və onun aid olduğu sinfin müəyyənləşdirilməsi üçün, uyğun olaraq, binar və çoxsinifli klassifikasiya metodları, həmçinin zərərli proqram nümunələrinin oxşar xüsusiyyətlər əsasında qruplaşdırılmasını təmin edən klasterləşdirmə metodu;
- Zərərli proqram təminatının binar fayllarının təsvir formasında təqdim edilməsinə və həmin təsvirlərin dərin öyrənmə alqoritmləri vasitəsilə emalına əsaslanan aşkarlama metodu;
- Şərti generativ rəqib şəbəkələrdən istifadə etməklə sintetik nümunələrin yaradılması, verilənlər çoxluğunda siniflərarası balanssızlığın azaldılması və zərərli proqram təminatının aşkarlanması modelinin dəqiqliyinin yüksəldilməsi metodu;
- Təklif olunan zərərli proqram aşkarlama metodlarının izahediləbilənliyinin, dayanıqlığının qiymətləndirilməsini, eləcə də dəyişən real mühit şəraitinə uyğunlaşma qabiliyyətinin yoxlanılmasını təmin edən eksperimental yanaşma.

Tədqiqatın elmi yeniliyi. Dissertasiya işi çərçivəsində aşağıdakı elmi nəticələr əldə edilmişdir:

- Statik və dinamik analizdən əldə edilən əlamətlərin diskretliyi, kəsilməzliyi və modelin xəttiliyi nəzərə alınmaqla informativ əlamətlərin seçilməsi üçün hibrid filtrasiya metodu işlənilmişdir.
- Zərərli proqramların aşkarlanması və növlərinin müəyyənləşdirilməsi üçün binar və çoxsinifli klassifikasiya metodları işlənilmişdir.
- Naməlum zərərli proqram qruplarının oxşar əlamətlər əsasında aşkarlanması üçün klasterləşdirmə metodu işlənilmişdir.
- Zərərli proqramların binar fayllarının təsvirə çevrilməsi və dərin öyrənmə əsasında klassifikasiyası üçün aşkarlama

metodu işlənilmişdir.

- Şərti generativ rəqib şəbəkələr əsasında sintetik nümunələrin yaradılması ilə sinif balanssızlığının azaldılması və modelin ümumiləşdirmə qabiliyyətinin artırılması metodu işlənilmişdir.
- Təklif olunan metodların real istismar mühitinə inteqrasiyası, dəyişən verilənlər şəraitində fəaliyyəti və dayanıqlığının qiymətləndirilməsi üçün model işlənilmiş və eksperimental yoxlanılmışdır.

Tədqiqatın nəzəri və praktiki əhəmiyyəti. *Dissertasiya işinin nəzəri əhəmiyyəti* zərərli proqramların aşkarlanmasında beynəlxalq təcrübənin təhlilinə, eləcə də maşın öyrənməsi modellərinin dəqiq və dayanıqlı işləməsini təmin edən üçmərhələli konseptual model və hibrid filtrasiya metodunun işlənilməsinə dayanaraq sahənin elmi-metodoloji bazasını zənginləşdirir. *İşin praktiki əhəmiyyəti* isə təklif olunan alqoritmlərin Windows, Android mühitlərində və mobil cihazlarda zərərli proqramların aşkarlanması, eləcə də təhlükəsizlik əməliyyat mərkəzlərində real vaxt rejimində təhdidlərin bloklanması və monitorinqi üçün birbaşa tətbiq edilə bilməsi ilə müəyyən olunur.

Aprobasiyası və tətbiqi. Dissertasiya işinin əsas nəzəri və praktiki nəticələri aşağıdakı siyahıda təqdim olunan bir sıra beynəlxalq və respublika səviyyəli konfranslarda məruzə edilmişdir:

- Tələbə və Gənc Tədqiqatçıların I (Bakı 2020), II (Bakı 2021), IV (Bakı 2023) Beynəlxalq Elmi konfransları;
- 15th (Baku 2021), 17th (Baku 2023) International Conference on Application of Information and Communication Technologies (AICT);
- XXV National Scientific Confernece (NASCO) of PhD student and young researchers, Baku, 2023;
- XVII International Scientific and Technical Conference "Optico-electronic devices in image recognition and image processing systems", Kursk, 2023;
- International Conference on Information Security:problems and prospects, Baku, 2022;
- National Supercomputer Forum (NSCF), (Pereslavl-Zalessky 2023);

- International Conference on Intelligent Systems and New Applications, Liverpool, 2023;
- VII Internatıonal Scientific Conference of Young Researchers, Baku, 2023;
- 5th International Conference on Problems of Cybernetics and Informatics (PCI), Baku, 2023;
- Heydər Əliyevin dövlət təhlükəsizliyi siyasəti: tarixi nailiyyətlər və müasir çağırışlar; mövzusunda respublika elmi-praktiki konfransı, Bakı, 2025.

Dissertasiya işinin yerinə yetirildiyi təşkilatın adı. Dissertasiya işi Azərbaycan Respublikası Elm və Təhsil Nazirliyinin İnformasiya Texnologiyaları İnstitutunda 2 saylı şöbədə yerinə yetirilmişdir.

Elmi nəşrlər. Dissertasiya işi üzrə 9-u elmi jurnalda, 14-ü isə respublika və beynəlxalq elmi konfransların materiallarında olmaqla 23 elmi iş dərc olunmuşdur. Elmi məqalələrdən 2-si Web of Science, 1-i Scopus, 1-i hər iki bazada indeksləşdirilən jurnallarda, 2-si digər beynəlxalq indeksləşdirmə sistemlərinə daxil olan jurnallarda, 3-ü isə AAK-ın tövsiyə etdiyi digər elmi jurnallarda dərc olunmuşdur. Konfrans materiallarından 3-ü xaricdə nəşr edilməklə 3-ü həm Scopus, həm də Web of Science bazalarında indeksləşdirilmişdir. Müəllifin Scopus bazasında h-indeksi 5, ona olan istinadların sayı isə 76-dır.

İşin strukturu və həcmi. Dissertasiya işi giriş, 4 fəsil, nəticə, 169 adda ədəbiyyat siyahısı, 40 cədvəl, 17 qrafik və 51 şəkildən ibarət olub, ümumi həcmi 175 səhifə təşkil edir.

İŞİN ÜMUMİ XARAKTERİSTİKASI

Girişdə dissertasiya mövzusu üzrə aktuallıq əsaslandırılmış, tədqiqat işinin məqsədi və vəzifələri müəyyən edilmişdir. Dissertasiya işi çərçivəsində əldə edilmiş nəticələrin elmi yeniliyi, nəzəri və praktiki əhəmiyyəti göstərilmişdir.

Birinci fəsildə zərərli proqram təminatının növləri, onların təbiətinin, eləcə də mövcud analiz metodlarının dərindən araşdırılması və tədqiq edilməsi vasitəsilə mövcud problemlərin aşkarlanması, eləcə də tədqiqatçılar tərəfindən zərərli proqram təminatının aşkarlanması istiqamətində aparılmış tədqiqatların analizi aparılmışdır. **Birinci fəslin birinci bölməsində** zərərli proqramların növləri, aşkarlanmaqdan yayınma üsullarının (metamorfizm, polimorfizm, paketləmə) iş prinsipi və bu üsullara uyğun əks-tədbirlər tədqiq edilmişdir [18]. **Birinci fəslin ikinci bölməsində** zərərli proqram təminatının mövcud analiz metodları tədqiq edilmiş, onların üstün və çatışmaz çəhətləri analiz edilmişdir. Statik və dinamik analiz üçün istifadə edilən üsul və vasitələr göstərilmiş, əsas tədqiqat obyekti kimi icraolunabilən faylların (PE fayllar) strukturu tədqiq edilmişdir [3]. **Birinci fəslin üçüncü bölməsində** zərərli proqram təminatının aşkarlanması üçün tətbiq edilmiş intellektual analiz metodlarının tədqiqi aparılmışdır. Tədqiqat üçün PRISMA 2020 metodikasından istifadə edilmiş, Scopus və Web of Science bazalarında indekslənən nəşrlər sistematik şəkildə təhlil edilmiş, üstünlükləri və çatışmazlıqları müəyyən olunmuşdur [3].

Aşağıda analiz nəticəsində müəyyən edilmiş boşluqlar verilmişdir:

1. Zərərli proqramın təsvirlər əsasında aşkarlanmasında əvvəlcədən öyrədilmiş modellərin təkmilləşdirilməsinin (*ing*. fine-tuning) hesablama yükü, çoxsaylı parametrləri, yüksək resurs tələbatı və əlamətlərin analizinin məhdudluğu kimi problemlər aşkarlanmışdır [4], [11], [20].

2. Generativ Rəqib Şəbəkələrin (GRŞ) zərərli proqram aşkarlanmasında tətbiqi ilə bağlı əsas problemlər [5], [12], [15]:

- *Verilənlərin mövcudluğu*: GRŞ-lərın effektiv öyrədilməsi üçün böyük və müxtəlif verilənlər bazası tələb olunur. Real praktikada zərərli proqram nümunələrini (xüsusilə sıfırıncı gün və polimorfik hücumlar) əldə etmək çətindir.

- *Sinif uyğunsuzluğu:* Zərərli proqram verilənlərində adətən normal nümunələrin sayı daha çox olur, bu da GRŞ-lərin nadir (zərərli) sinif nümunələrini generasiya etməsini çətinləşdirir.
- *Resurs tələbi*: GRŞ-lərın öyrədilməsi yüksək hesablama və vaxt tələb edir.
- *Qiymətləndirmə metrikaları*: Generasiya edilmiş zərərli nümunələrin təsirini ölçmək üçün uyğun metrikaların olmaması problemi mövcuddur.

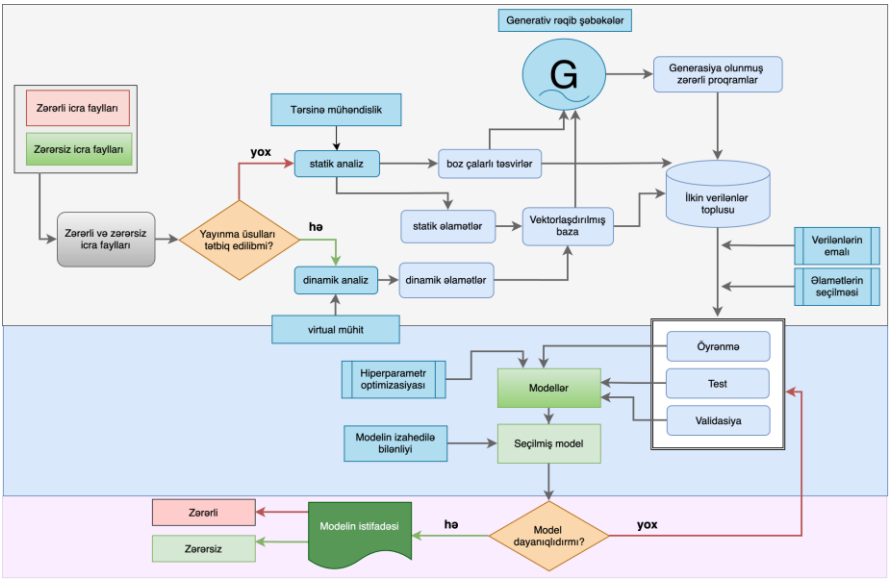
3. Zərərli proqram aşkarlama modellərinin izahediləbilənliyi və real mühitdə tətbiqi üzrə müəyyən edilmiş çatışmazlıqlar [17], [22]:

- Aşkarlama üçün təklif edilmiş modellərin yalnız eyni növ zərərli proqramlardan ibarət bazaya etibar etməsi modelin dayanıqlığını şübhə altında saxlayır.
- Aşkarlama üçün təklif edilmiş modellərin real zaman mühitində test edilməsi hallarına analiz edilmiş tədqiqat işlərində rast gəlinməmişdir.
- Bir çox tədqiqatda zərərli proqram təminatının aşkarlanması prosessində modelin dəqiqliyini süni şəkildə artırmaq üçün verilənlər toplusunda zərərli və zərərsiz fayllar məqsədli şəkildə bərabər paylanmaqla götürülmüş və ya az sayda verilənlərin dublikatını yaratmaqla çoxaldılaraq bərabərləşdirilmişdir. Bu hallar modelin dəqiqliyini artırsa da test mühitində dəqiqliyin kəskin azalması ilə nəticələnə bilər və analiz edilmiş tədqiqatlarda bu, yoxlanılmamışdır.
- Mövcud modellərin əksəriyyətində izahediləbilənlik problemi qalmaqdadır. Yəni modelin hansı əlamətlərə əsaslanaraq qərar qəbul etməsi aydın şəkildə izah olunmur. Bu isə istifadəçilərin sistemə etimadını azaldır, yanlış pozitiv və yanlış neqativ nəticələrin səbəblərinin müəyyənləşdirilməsini çətinləşdirir və real mühitdə tətbiq imkanlarını məhdudlaşdırır.

Birinci fəslin dördüncü bəndində zərərli proqram təminatının aşkarlanması üçün intellektual analiz metodlarının və alqoritmlərinin işlənməsi üzrə qarşıya qoyulmuş tədqiqat məsələləri əvvəlki bənddə verilmiş mövcud problemlər arasından prioritetliyə əsasən seçilmiş və tədqiqat məsələləri formalaşdırılmışdır.

İkinci fəsil zərərli proqram təminatının aşkarlanması sisteminin

konseptual modelinin işlənilməsinə həsr edilmişdir. **İkinci fəslin birinci bəndində** şəkil 1-də verilmiş zərərli proqram təminatının aşkarlanması sisteminin ardıcıl, üç mərhələdən ibarət konseptual modeli təklif edilmişdir [21]. Birinci mərhələ zərərli proqram təminatının aşkarlanması modelinin qurulması anına qədər olan hissəni əhatə edir. İkinci mərhələ modelin qurulması, test edilməsi, yaxşılaşdırılması mərhələsidir. Burada daha dəqiq və dayanıqlı metodun seçilməsi üçün, eyni zamanda çoxlu sayda əlamətlər arasından daha uyğun əlamətlərin seçilməsinə dair metodlar işlənmişdir.



Şəkil 1. Zərərli proqram təminatının aşkarlanması üçün təklif edilmiş konseptual model

Üçüncü mərhələ ikinci mərhələdən əldə edilmiş yekun metodun izaholunma qabiliyyətinin göstərilməsi və modelin real mühitdə dayanıqlığının yoxlanılması və monitorinqi məsələlərinə həsr edilmişdir.

Təklif edilmiş konseptual modeldən göründüyü kimi, ilk mərhələdə zərərli proqram aşkarlama modelinin qurulmasına hazırlıq üçün müəyyən əlamətlər fəzasından ibarət verilənlər bazası formalaşdırılır. Əlamətlər statik və dinamik yollarla əldə edilə bilər və hansı növ

əlamətlərin seçilməsi probleminə həll olaraq **ikinci fəslin ikinci bəndində** icra olunan fayllardan əlamətlərin seçilməsi üçün əlamətlərin diskretliyi və kəsilməzliyi, modelin xəttiliyi kimi faktorlarını nəzərə alan iki mərhələli metod işlənmişdir [23]. Təklif edilmiş metodun tətbiqi nəticəsində modelin qurulması və istismarı mərhələsində yüksək hesablama resursuna olan tələbat əhəmiyyətli dərəcədə azaldılmış olacaqdır.

Birinci mərhələ. Birinci mərhələ modeldən asılı olmadan həyata keçirilən ilk mərhələdir. Bu mərhələdə verilənlər toplusunda sabit və ya dəyişməz əlamətlər (variasiyası sıfra bərabər olan), nisbətən sabit əlamətlər, dublikat əlamətlər müəyyən edilir və kənarlaşdırılır. Bu əlamətlər verilənlər toplusunda variasiya yaratmadığı üçün modellərin proqnozlaşdırma gücünə töhfə vermir. Kəsilməz əlamətlər üçün dispersiyanın hesablanması formul 1-də verilmişdir.

$$V(X_i) = \frac{1}{n}\sum_{j=1}^{n}(x_{ij} - \hat{x}_{ij})^2, \quad (1)$$

burada $V(X_i) - X_i$ dəyişəninin dispersiyası, $x_{ij} - X_i$-nin j-ci müşahidə dəyəri, $\hat{x}_{ij} - X_i$-nin orta dəyəri, n - ümumi müşahidə sayını bildirir. Əldə edilmiş dispersiya qiyməti əvvəlcədən təyin edilmiş sərhəd qiyməti ilə müqayisə edilir.

Diskret əlamətlər üçün dominant kateqoriyanın faizi hesablanır və əvvəlcədən təyin edilmiş sərhəd qiyməti ilə müqayisə edilərək əlamətin çıxarılıb-çıxarılmayacağına qərar verilir. Diskret əlamətlər üçün dominant kateqoriyanın faizinin hesablanması formul 2-də verilmişdir.

$$D(X_i) = \max\left(\frac{count(c_k)}{n}\right), \forall c_k \in X_i \quad (2)$$

burada $D(X_i)$ - dominant kateqoriyanın faizi, $count(c_k) - c_k$ - kateqoriyasının müşahidə sayı, n - ümumi müşahidə sayını bildirir. Əldə edilmiş dominant kateqoriya faizi əvvəlcədən təyin edilmiş sərhəd qiyməti ilə müqayisə edilir.

Kəsilməz əlamətlərdə dublikat əlamətlərin çıxarılması üçün formul 3-də verilmiş Pirson korrelyasiyasından istifadə edilir:

$$r(X_i, X_j) = \frac{\sum(X_i - \overline{X}_\iota)(X_j - \overline{X}_J)}{\sqrt{\sum(X_i - \overline{X}_\iota)^2 \sum(X_j - \overline{X}_J)^2}}, \quad (3)$$

burada $r(X_i, X_j)$ - korrelyasiya əmsalı, $\overline{X}_\iota - X_i$-nin, $\overline{X}_J - X_j$ -in orta qiymətidir. Seçim şərti üçün əldə edilmiş korrelyasiya qiyməti təyin edilmiş sərhəd qiyməti ilə müqayisə əsasında müəyyən edilir[4].

Diskret əlamətlərdə dublikat əlamətlərin çıxarılması üçün formul 4-də verilmiş Spirman ranqlaşdırma korrelyasiyasından istifadə edilir[5].

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)}, \quad (4)$$

burada d_i - hər bir müşahidə üçün iki dəyişənin ranqları arasındakı fərq, n - müşahidələrin ümumi sayıdır. Əgər ranqlar arasında heç bir korrelyasiya yoxdursa, ρ-in qiyməti 0-a bərabər hesab edilir.

İkinci mərhələ. Bu mərhələdə əlamətlərin seçilməsi zamanı hansı modelin tətbiq edildiyi nəzərə alınır və seçilmə qaydası ona uyğun olaraq təyin edilir. Təklif edilmiş üsulda əlamətlərin həm individual, həm qarşılıqlı təsiri nəzərə alınmaqla dəqiq əlamət seçilməsi reallaşdırılır. Formul 5-də əlamətlərin təsir gücünün individual olaraq hesablanması qaydası verilmişdir.

$$T_i(X_i) = \frac{\sum E(X_i)}{S}, \quad (5)$$

burada $T_i(X_i)$ - əlamətlərin individual təsir gücü, E - entropiya, S isə ümumi bölünmələrin sayını bildirir.

Əlamətlərin qarşılıqlı təsirinin nəzərə alınması üçün birlikdə istifadə edilən əlamətlərin bölünmə sayına uyğun olaraq entropiyalar hesablanır və cəmlənir. Formul 6-da əlamətlərin birgə istifadəsinin təsir gücünün hesablanması qaydası verilmişdir.

$$T_g(X_i, X_j) = \sum E(X_i, X_j), \quad (6)$$

[4] Ge, G., Zhang, J. Feature selection methods and predictive models in CT lung cancer radiomics // Journal of Applied Clinical Medical Physics, – 2023. 24(1), – Article e13869.

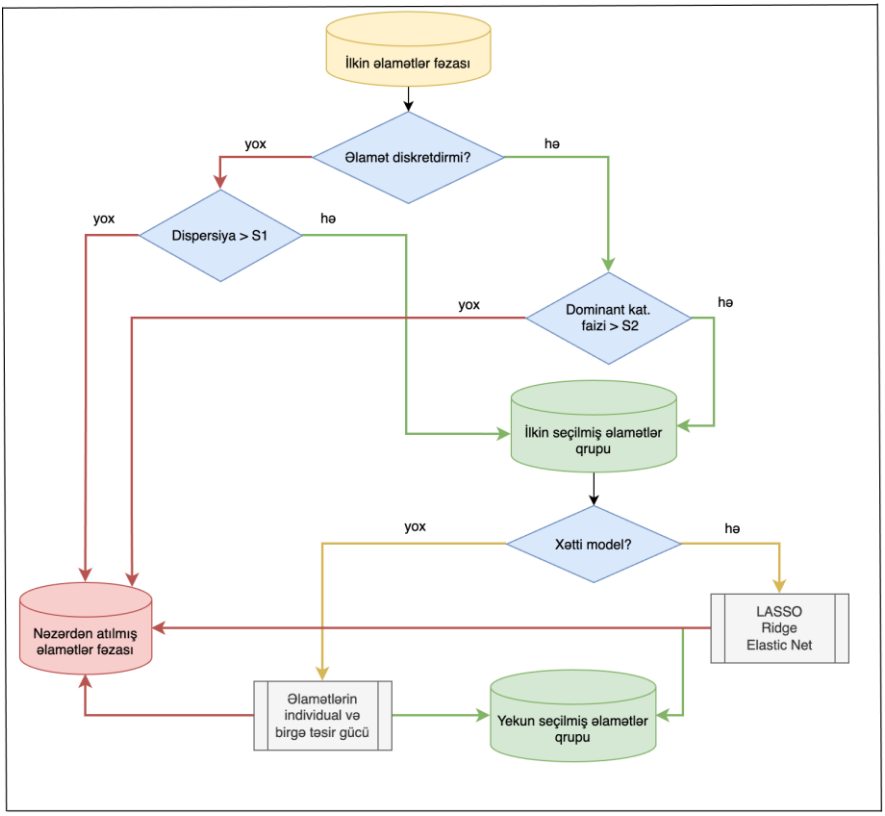
[5] Kim, D., Heo, T.-Y. Anomaly Detection with Feature Extraction Based on Machine Learning Using Hydraulic System IoT Sensor Data // Sensors, – 2022. 22(7), – p. 1-24.

burada $T_g(X_i, X_j)$ - əlamətlərin birgə təsir gücünü bildirir.

Yekun olaraq, əlamətlərin həm individual, həm birgə təsirinin nəzərə alındığı ümumi qayda formul 7-də verilmişdir.

$$T_y(X_i) = \alpha * T_i(X_i) + \beta * \sum_{i \neq j} T_g(X_i, X_j), \quad \alpha + \beta = 1 \qquad (7)$$

burada $T_y(X_i)$ - əlamətin yekun təsir gücü, α və β isə çəkilərdir. Çəkilər istifadəçi tərəfindən təyin edilən və 0 ilə 1 arasında dəyişən qiymətlərdir. Şəkil 2-də təklif edilmiş yanaşmanın ümumi blok-sxemi göstərilmişdir [23].



Şəkil 2. Əlamətlərin seçilməsi üçün təklif edilmiş yanaşmanın blok-sxemi

Zərərli proqram təminatının aşkarlanması üçün təklif edilən metod, xüsusilə ağac əsaslı modellərdən istifadə edilərək əlamətlərin

seçilməsinin daha dəqiq və effektiv şəkildə həyata keçirilməsini hədəfləyir.

İkinci fəslin üçüncü bəndində zərərli proqram təminatının aşkarlanması modellərinin izahediləbilənliyi və real mühitlə inteqrasiyası metodlarının işlənməsinə baxılmışdır. SHAP və LIME kimi mövcud izahediləbilənlik metodlarının çatışmazlıqlarını nəzərə alaraq, zərərli proqramların aşkarlanması modelləri üçün izahetmə sabitliyini ölçən riyazi qiymətləndirmə metodu təklif edilmişdir. Təklif olunan yanaşmanın əsas ideyası ondan ibarətdir ki, eyni sinfə aid olan və oxşar xüsusiyyətlərə malik nümunələr üçün modelin verdiyi izahlar da statistik baxımdan bir-birinə yaxın olmalıdır. Əgər izahlar yüksək variasiya göstərirsə, bu modelin qərar mexanizminin qeyri-stabil olduğunu və bəzi əlamətlərə təsadüfi şəkildə həddindən artıq reaksiya verdiyini göstərir.

Bu məqsədlə əvvəlcə eyni sinfə aid olan K ədəd nümunə seçilir və hər bir nümunə üçün SHAP və ya LIME metodu vasitəsilə formul 8-də verilmiş əlamət təsir vektorları hesablanır:

$$\Phi^{(k)} = \left(\phi_1^{(k)}, \phi_2^{(k)}, \dots, \phi_n^{(k)}\right), k = 1,2, \dots, K, \quad (8)$$

burada n – modeldə istifadə olunan əlamətlərin sayını, $\phi_i^{(k)}$ isə k-cı nümunə üçün i-ci əlamətin izahedici təsir qiymətini göstərir. Hər bir əlamət üzrə izahetmə dəyişkənliyi formul 9-da verilmiş qaydada müəyyən edilir:

$$\mathrm{Var}_i = \mathrm{Var}\left(\phi_i^{(1)}, \phi_i^{(2)}, \dots, \phi_i^{(K)}\right) \quad (9)$$

Bu dəyişkənlik əsasında əlamət üzrə izahetmə sabitliyi göstəricisi formul 10-dakı kimi təklif olunur:

$$\mathrm{Stab_i} = 1 - \frac{\mathrm{Var}_i}{\max\limits_{j=1,\dots,n} \mathrm{Var}_j}, \quad (10)$$

burada $\mathrm{Stab}_i \in [0,1]$ intervalında qiymət alır və i-ci əlamətin izahının nə dərəcədə stabil olduğunu ifadə edir. Qiymətin 1-ə yaxın olması əlamətin model qərarına verdiyi töhfənin nümunələr arasında sabit olduğunu, 0-a yaxın olması isə izahın qeyri-stabil xarakter daşıdığını göstərir.

Model üzrə ümumi izahetmə sabitliyi isə aşağıda verilmiş qaydada müəyyən edilir:

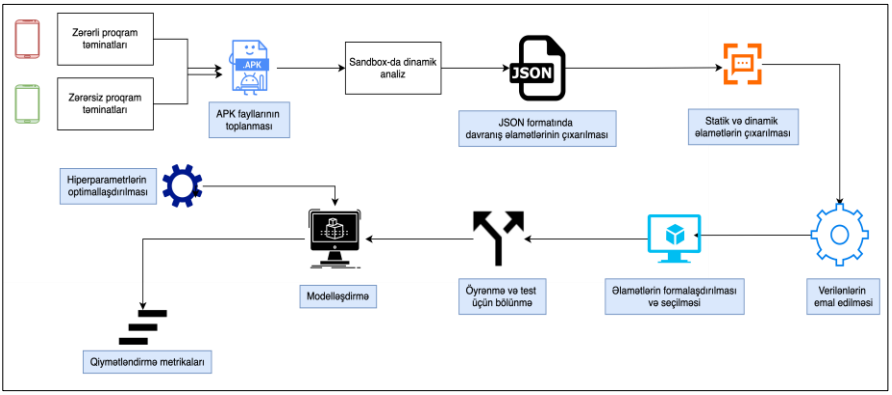
$$\text{Stab}(f) = \frac{1}{n}\sum_{i=1}^{n} \text{Stab}_i \quad (11)$$

Bu göstərici modelin izahetmə baxımından ümumi etibarlılıq səviyyəsini xarakterizə edir və müxtəlif aşkarlama modellərinin (məsələn, Random Forest, XGBoost, CNN, LSTM və s.) izahetmə keyfiyyətinin müqayisəli təhlili üçün istifadə oluna bilər.

İkinci fəslin üçüncü bəndində aşkarlama modellərinin real mühitdə istifadəsi zamanı dayanıqlığın yoxlanması məqsədilə müxtəlif məsələlərə baxılmışdır [14]. Bunlardan birincisi verilənlərdə dəyişikliyin yoxlanılması üçün PSI metodunun istifadəsi, ikincisi modelin göstərdiyi performansda yayınmanın izlənməsidir. Beləliklə, qurulacaq sistem yalnız laboratoriya şəraitində yüksək dəqiqlik göstərməklə kifayətlənmir, həm də real mühitdə davamlılıq, etibarlılıq və şəffaflıq baxımından qiymətləndirilir.

Üçüncü fəsildə Windows və Android əməliyyat sistemində lokal mühitdə rast gəlinən zərərli proqramlar üzərində həm binar, həm də çoxsinifli klassifikasiya üsulları, klasterləşdirmə üsulları, təsvirlər əsasında klassifikasiya üsulları, Generativ Rəqib Şəbəkə əsaslı üsullar, avtoenkoder metodlarının tətbiqi əsasında yeni modellər qurulmuşdur. Həm Windows, həm də Android əməliyyat sistemi proqramları üçün eksperimental tədqiqatlar ikinci fəsildə təklif edilmiş konseptual modelin birinci və ikinci mərhələsinə müvafiq olaraq aparılmışdır. **Üçüncü fəslin birinci bəndində** zərərli proqram təminatının aşkarlanması üçün maşın öyrənmə metodlarının (binar klassifikasiya, çoxsinifli klassifikasiya, klasterləşdirmə) işlənməsinə baxılmışdır [1], [2], [6-8], [10], [13], [19]. Bu işlərdən birincisi zərərli proqramlardan çıxarılmış opkodlar (əməliyyat kodları, *ing*. operational code) əsasında yüksək dəqiqliyə malik maşın öyrənmə modelinin qurulması olmuşdur [13]. Nəticə olaraq müşahidə edilmişdir ki, opkod tezliyi əsasında hazırlanmış verilənlər toplusu üzərində ağac alqoritmləri nisbətən yuxarı nəticə göstərsə də bəzi çatışmazlıqlar mövcuddur. Belə ki, əlamətlərin əhəmiyyətliliyi əsasında yüksək rastlama tezlikli "push", "move" kimi opkodların həm zərərli, həm də zərərsiz proqramlarda eyni əhəmiyyətliliyi göstərməsi bu opkodların ayırdedici xüsusiyyəti olmasını şübhə altına salır. Bu səbəbdən digər əlamətlərin tədqiq olunması zərurəti yaranır [13]. İkinci məsələdə, xüsusi hal olaraq, zərərli proqram trafikinin aşkarlanması məqsədilə “MTA-KDD19”

zərərli proqram trafiki verilənlər bazasından istifadə edərək yüksək dəqiqliyə malik maşın öyrənmə modeli qurulmuşdur [7]. Əsas tədqiqat obyektinə daxil olan Android əməliyyat sistemində "CICMalDroid" və "Drebin" verilənlər toplusundan istifadə etməklə zərərli proqram təminatının aşkarlanmasının təkcə binar halına deyil, çoxsinifli halına da bu bənddə baxılmışdır [6], [8], [10]. Şəkil 3-də təklif edilmiş ümumi sxem əks etdirilmişdir.



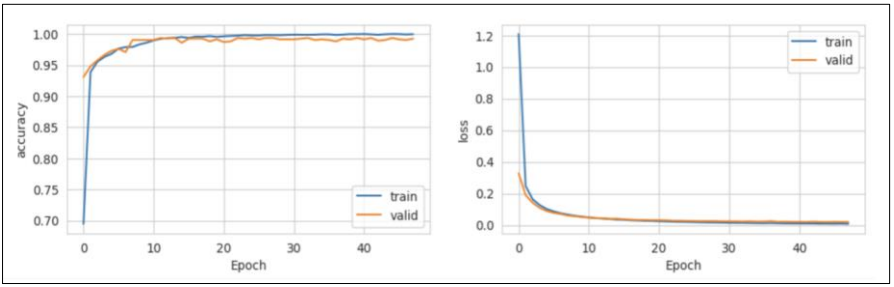
Şəkil 3. Zərərli mobil tətbiqlərin aşkarlanması üçün təklif edilmiş sxem

Nişanlanmamış proqram nümunələrinin oxşar əlamətlər əsasında qruplaşdırılması məqsədilə k-modes klasterləşdirmə alqoritmi tətbiq edilmişdir. Klasterlərin optimal sayının müəyyənləşdirilməsi üçün müxtəlif k qiymətləri üzrə alınmış nəticələr qiymətləndirilmiş və seçilmiş verilənlər toplusunda k=2 qiyməti daha əlverişli nəticə göstərmişdir [2].Tədqiqat zamanı aşağıdakı nəticələrə gəlinmişdir.

- Siqnatura əsaslı metodlar yeni və dəyişkən siqnaturaya malik zərərli proqramları aşkarlaya bilmir.
- Dinamik analiz çox effektiv olsa da, böyük miqyaslı sistemlərdə real vaxtda tətbiqi mümkün deyil.
- Hibrid yanaşmaların qurulması çətinliklər yaradır və çox vaxt sistemlərarası uyğunluq problemləri yaşanır.
- Zərərli proqram təminatının getdikcə daha mürəkkəb olması, onların davranışlarını gizlətmək üçün şifrləmə və maskalama üsullarından istifadə etməsi, ənənəvi aşkarlama üsullarını arxa plana keçirir [2].

Üçüncü fəslin ikinci bəndində zərərli proqram təminatının dərin təlim əsasında aşkarlanması metodunun işlənilməsinə baxılmışdır. Mövcud verilənlər toplusu ilə yüksək nəticələr əldə edilsə də, dəqiqliyin əlamətlərin keyfiyyətindən asılı qalması, yeni növ zərərli proqramları aşkarlamaq məhdudiyyətləri göstərir ki, aşkarlamanın etibarlığını artırmaq üçün təsvirlər əsasında aşkarlanma metodlarının işlənməsi və empirik sınaqdan keçirilməsi elmi və praktiki baxımdan zəruridir. Davranış analizində ən çox istifadə olunan iki əlamət ailəsi — sistem çağırışları və API çağırışlarıdır. Dərin öyrənmə modelləri bu əlamətlər əsasında zərərli proqramı müəyyənləşdirmə və təsnifetmədə ümidverici nəticələr göstərsə də, aşkarlamadan yayınmağa yönəlmiş zərərli proqramlara qarşı dayanıqlıq məsələsi, eləcə də dərin öyrənmə modellərinin yüksək hesablama/yaddaş tələbləri nəzərə alınmalıdır [9].

Bu istiqamətə birinci yanaşmada zərərli proqram nümunələri təsvirlərə çevrilərək dərin öyrənmə metodları vasitəsilə zərərli proqram təminatının klassifikasiyası [4], [11], [20], ikinci yanaşma isə təsvirlər üzərində avtoenkoder metodunun tətbiqi vasitəsilə anomaliyaların aşkarlanması məsələsinə gətirməkdir [16]. Təsvirlər əsasında aşkarlama üsulunda PE fayllar əvvəlcə binar kodlara çevrilir və təsvirlərin alınması üçün standart ölçüyə gətirilir. Müxtəlif çevirmə üsulları vasitəsilə standart ölçülü boz çalarlı təsvirlər əldə edilir [16]. "Malimg" verilənlər toplusu üzərində 13 transfer öyrənmə (*ing*. transfer learning) metodları yoxlanılmış və MobileNet modeli ilə yüksək dəqiqlik və dayanıqlığa malik model inkişaf etdirilmişdir. Şəkil 4-də period artdıqca öyrənmə və validasiya verilənləri üzrə dəqiqliyin və itkinin ədədi qiymətinin dəyişməsi verilmişdir.



Şəkil 4. MobileNet üzrə nəticələr

Qrafikdən aydın görünür ki, öyrənmə və validasiya verilənləri üzrə alınmış nəticələr həm yüksək, həm də bir-birindən az fərqlənmişdir. Bu isə modelin daha dayanıqlı olduğunun göstəricisidir. Cədvəl 1-də təklif edilmiş metod ilə digər müəlliflər tərəfindən eyni verilənlər toplusundan istifadə etməklə alınmış nəticələrin müqayisəsi aparılmışdır.

Cədvəl 1. "Malimg" verilənlər toplusu üzərində aparılmış zərərli proqram aşkarlama metodlarının müqayisəli analizi

Müəllif	İl	Təklif olunan metod	F1 ölçü (%)
Hai T. və b.	2023	Inception V3[6]	74.29
Kim J. və b.	2023	Diqqət mexanizminə əsaslanan Cross-modal CNN[7]	97.6
Ahmed, I.T. və b.	2024	Seqmentasiyaya əsaslanan fraktal tekstura analizi və KNN[8]	70.14
Təklif olunan metod	**2024**	**MobileNet**	**98.9**

Təsvirlər əsasında əldə olunan yüksək göstəricilərə baxmayaraq, nəticələrin sinif balanssızlığına, verilənlərin müxtəlifliyinə və paylanma dəyişmələrinə həssaslığı kimi problemlər də qalır. Şəkillərə çevirmə mərhələsi binar fayldakı bəzi semantik siqnalları itirə, az təmsil olunan siniflərdə ümumiləşdirməni zəiflədə və yeni yayınma üsullarına qarşı dayanıqlığı azalda bilər. Bundan əlavə, sinifləndirmə xərci və sinif keyfiyyətinə asılılıq praktiki tətbiqlərdə məhdudiyyət

[6] Hai, T.H. A Proposed New Endpoint Detection and Response With Image-Based Malware Detection System / T.H. Hai, V. Van Thieu, T.T. Duong [et al.] // IEEE Access, – 2023. 11, – p. 122859–122875.

[7] Kim, J., Paik, J.Y., Cho, E.S. Attention-Based Cross-Modal CNN Using Non-Disassembled Files for Malware Classification // IEEE Access, – 2023. 11, – p. 22889–22903.

[8] Ahmed, I.T., Hammad, B.T., Jamil, N. A Comparative Performance Analysis of Malware Detection Algorithms Based on Various Texture Features and Classifiers // IEEE Access, – 2024. 12, – p. 11500–11519.

yaradır. Bu səbəbdən zərərli proqram təminatının aşkarlanmasını anomaliya aşkarlanması kimi baxaraq avtoenkoderlərdən istifadə edilmişdir [16]. Təklif olunan yanaşmada girişlər ya boz-çalarlı zərərli proqram təsvirləri, ya da icra əlamət vektorları kimi formalaşdırılır, enkoder–dekoder arxitekturası (məsələn, konvolyusiya avtoenkoderi) və ya struktur oxşarlıq tipli itkilərlə öyrədilir. Aşkarlama üçün sərhəd dəyəri validasiya verilənləri üzərində ROC analizi və ya F1-ölçü maksimumu ilə kalibrasiya edilir. Bu yanaşma iki üstünlük verir: (i) sinif asılılığını azaldır və yeni ("zero-day") ailələrə qarşı həssaslığı artırır; (ii) mövcud təsvir-əsaslı nəzarətli modelləri hibrid sxemdə tamamlayaraq əvvəlcədən filtrləmə/siqnal gücləndirmə mərhələsi kimi işləyə bilir. Klassifikasiya performansını maksimum edən optimal həddi tapmaq üçün bir neçə hədd dəyəri sınaqdan keçirildi. Bir neçə testdən sonra 0.95 hədd dəyərinin ən yaxşı nəticələri verdiyi müəyyənləşdirildi. Belə ki, daxil olan təsvirin bərpa xətası 0.95-dən aşağıdırsa, nümunə zərərli proqram kimi, 0.95-dən yuxarıdırsa, zərərsiz kimi təsnif edilir. Model test verilənləri üzərində qiymətləndirildi, xəta matrisi və dəqiqlik, həssaslıq və F1-ölçü daxil olmaqla müxtəlif performans metrikaları hesablandı. Xəta matrisinə əsasən, modelin ümumi doğruluğu 62.3%-dir. Zərərli kimi proqnozlaşdırılan nümunələr içində həqiqətən zərərli olanların payını ölçən həssaslıq metrikası 58.1% olmuşdur. Modelin faktiki zərərli nümunələri aşkar etmə qabiliyyətini göstərən tamlıq isə yüksək - 91.9% qiyməti almışdır [16]. Zərərli proqram təminatının aşkarlanması kimi məsələlərdə tamlıq xüsusilə vacibdir, çünki zərərli nümunələrin gözardı edilməsi ciddi təhlükəsizlik risklərinə səbəb ola bilər. **Üçüncü fəslin üçüncü bəndində** verilənlər bazasındakı balanssızlıq problemini həll etmək üçün Şərti Generativ Rəqib Şəbəkə əsaslı model təklif edilmişdir [15]. Şərti GRŞ-lərdə generasiya prosesi yalnız təsadüfi küydən (z) asılı qalmır, həm də əlavə şərt (y) – sinif, atribut və ya digər məlumat əsasında yönləndirilir. Bu yanaşma generatorun şərtə uyğun nümunələr yaratmasına, diskriminatorun isə həm nümunənin real/saxta olmasını, həm də şərtlə uyğunluğunu qiymətləndirməsinə imkan verir. Beləliklə, Şərti GRŞ-lər nəzarətli generasiya problemlərində (məsələn, sinif balanssızlığının aradan qaldırılması, təsvirin digər təsvirə çevrilməsi, atributa görə təsvir

sintezi) geniş tətbiq edilir[9]. Şərti GRŞ-in məqsəd funksiyası 12-ci formulda təsvir edilmişdir [15]:

$$min_G max_D V(D,G) = E_{(x,y)\sim p_{data}(x,y)}[logD(x|y)] + \\ + E_{z\sim p_z(z),y\sim p(y)}\big[log\big(1 - D(G(z|y)|y)\big)\big] \qquad (12)$$

burada $y -$ şərt/etiket (məsələn, sinif və ya atribut), $D(x|y)$ - diskriminator, nümunənin real olub-olmamasını və şərtlə uyğunluğunu qiymətləndirir, $G(z|y) -$ generator, şərtə uyğun saxta nümunələr yaradır, $p_{data}(x,y) -$ real verilənlərin birgə paylanması, $p_z(z) -$ generatorun giriş səsi (məsələn, normal və ya binomial paylanma) olduğunu bildirir. Cədvəl 2-də generasiya edilməmişdən öncəki baza üzərində modellərin müqayisəsi, cədvəl 3-də isə Şərti GRŞ tətbiq edildikdən sonra balanslaşdırılmış baza üzərində modellərin müqayisəsi verilmiş və nəticələrin yaxşılaşdığı müşahidə edilmişdir.

Dördüncü fəsildə müxtəlif açıq verilənlər topluları əsasında zərərli proqram təminatının aşkarlanması modellərinin qərarlarının izahı və zamanla dəyişən verilənlər şəraitində dayanıqlığının qiymətləndirilməsi həyata keçirilmişdir. Əldə edilmiş nəticələr modellərin real istismar mühitində tətbiqi imkanlarının qiymətləndirilməsinə xidmət edir.

Cədvəl 2. Balanssız verilənlər toplusunda modellərin müqayisəsi

Modellər	Dəqiqlik	Həssaslıq	Tamlıq	F1-ölçü
LightGBM	0.9153	0.8958	0.8643	0.8798
Təsadüfi meşə	0.9189	0.9096	0.8593	0.8837
KNN	0.8559	0.8083	0.7839	0.7959
Loqistik reqressiya	0.7532	0.7460	0.4724	0.5785

[9] Imamverdiyev, Y.N., Musayeva, F.I. Analysis of generative adversarial networks // Problems of Information Technology, – Baku: – 2022, – 13 (1), – p. 20–27.

Cədvəl 3. Balanslı verilənlər toplusunda modellərin müqayisəsi

Modellər	Dəqiqlik	Həssaslıq	Tamlıq	F1-ölçü
LightGBM	***0.9568***	***0.9619***	***0.9541***	***0.9579***
Təsadüfi meşə	0.9513	0.9666	0.9378	0.9520
KNN	0.8955	0.9041	0.8919	0.8980
Loqistik reqressiya	0.8078	0.8816	0.7243	0.7953

Zərərli proqram təminatının aşkarlanması üçün qurulmuş modelin real mühitdə praktiki etibarlılığı və dayanıqlığı iki tamamlayıcı istiqamətdə əsaslandırılır:

(i) modelin izahediləbilənliyi [17], [19], [22];

(ii) real mühitdə dayanıqlığın sistemli monitorinqi [14].

Əsas məqsəd seçilmiş yekun modelin real mühitlərdə tətbiq edilə bilən, etibarlı və davamlı nəticələr əldə edə biləcəyini sübut etməkdir. Yekun modelin tətbiqi yalnız nəzəri və eksperimental şəraitdə uğurlu olmasını deyil, həmçinin praktik mühitlərdə də yaxşı performans göstərməsini təmin etməyi hədəfləyir. **Dördüncü fəslin birinci bəndində** üç fərqli verilənlər toplusundan istifadə etməklə zərərli proqram təminatının aşkarlanması modellərinin izahediləbilənliyinə real mühitdə praktiki olaraq baxılmışdır [17], [19], [22]. Bunun üçün, ilk növbədə, fərqli verilənlər toplusu üzrə müxtəlif maşın öyrənmə modelləri qurulmuş, öyrənməyə sərf olunan vaxt da nəzərə alınmaqla müxtəlif metriklərlə müqayisə edilmişdir. Aparılmış eksperimentlərin ümumi nəticələri göstərir ki, "CICMalDroid2020", "CICMalMem2022" və "Irfan Yousuf" verilənlər toplusu üzərində sınaqdan keçirilən modellər arasında LightGBM həm çoxsinifli, həm də binar təsnifat ssenarilərində ən stabil və yüksək performanslı yanaşma kimi seçilir:

- "CICMalDroid2020": F1=0.93 və ROC-AUC=0.99 [22]
- "CICMalMem2022": F1=0.81 və ROC-AUC=0.95 [19]
- "Irfan Yousuf": F1=0.93 və ROC-AUC=0.99 [17].

Real mühitdə izahediləbilənlik baxımından SHAP və LIME

analizləri modelin qərarlarının "məntiqi" əsasını açıq göstərmişdir: Android mühitində siniflər üzrə qərarlara əsasən cihaz identifikatorları və şəxsi məlumatlara çıxış (getDeviceId, getSubscriberId, ACCESS_PERSONAL_INFO), şəbəkə aktivliyi (NETWORK_ACCESS), əlçatanlıq/istifadəçi interfeysi manipulyasiyası (sendAccessibilityEvent) və fayl/sistem çağırışları (pread64 və s.) əsas təsiredici əlamətlər kimi çıxış etmiş; yaddaş verilənləri əsasında (CICMalMem2022) proses/xidmət və emal tipli əlamətlər sinifləri ayırd etmiş; Windows PE mühitində isə başlıq, bölmə və DLL/API ilə bağlı struktur göstəricilər (məs., TimeDateStamp, DLLCharacteristics, CheckSum və s.) zərərli proqramların fərqləndirilməsində əsas rol oynamışdır. Ümumilikdə, nəticələr LightGBM-in təkcə yüksək dəqiqlik deyil, həm də SHAP/LIME vasitəsilə qərarların izah oluna bilməsi sayəsində real mühitdə tətbiq üçün daha uyğun və etibarlı model olduğunu təsdiqləyir.

Dördüncü fəslin ikinci bəndində zərərli proqram təminatının aşkarlanması modellərinin dayanıqlığının real mühitdə qiymətləndirilməsi məsələsinə praktiki olaraq baxılmışdır [14]. Eksperiment üçün 2010-cu ilin avqust ayından 2012-ci ilin oktyabr ayına qədər qeydə alınmış zərərli proqramlardan ibarət Drebin verilənlər toplusundan istifadə edərək müxtəlif zaman intervallarında modelin dayanıqlığı təklif edilmiş üsullarla yoxlanılmışdır. Modelin qurulması üçün ilk növbədə verilənlər toplusunda ilk iki ilin verilənləri istifadə edilmiş, sonuncu il isə dayanıqlığın yoxlanılması məqsədilə saxlanılmışdır. Öyrənmə məqsədilə müxtəlif növ alqoritmlər - loqistik reqressiya və bir çox ağac yönümlü alqoritmlər istifadə edilmişdir. LightGBM alqoritmi F1 = 0.98 olmaqla digər metodlar arasından daha yüksək performans və dayanıqlıq göstərmişdir. Növbəti mərhələdə LightGBM alqoritminin stabilliyini yoxlamaq məqsədilə yenidən eksperiment üçün ayırılan sonuncu il məlumatları üzrə öyrənmə prosesi təkrarlanmışdır. Hər iki modelin əlamətlərinin əhəmiyyətlilik dərəcələri müqayisə edilmiş və bu əlamətlər arasında fərq müşahidə edilməmişdir. Əlavə olaraq, populyasiyanın stabilliyini yoxlamaq məqsədilə birinci və ikinci eskperimentdə istifadə olunan verilənlər toplusu PSI metodu ilə qiymətləndirilmiş və cədvəl 4-də yayınma

qiyməti yalnız 0.01 və ondan yüksək olan əlamətlər və yayınma qiymətləri əks edilmişdir.

Cədvəl 4. PSI göstəriciləri

Əlamətlər	Yayınma qiyməti
Landroid.content.Context.unregisterReceiver	0.02
Ljava.lang.Class.getResource	0.01
Ljava.net.URLDecoder	0.01
WRITE_HISTORY_BOOKMARKS	0.01
transact	0.01
TelephonyManager.getLine1Number	0.01
SEND_SMS	0.01
onServiceConnected	0.01
Ljava.lang.Class.getMethods	0.01
RECEIVE_SMS	0.01

Cədvəldən göründüyü kimi, verilənlər toplusunda kiçik dəyişikliklər modelin performansına ciddi təsir etmir və model stabil qalır. Aparılan müqayisəli eksperimentlər və yayınmanın monitorinqi sübut edir ki, yekun həll yeni verilənlərə qarşı dayanıqlıdır və yanlış qərar riskini minimuma endirir.

Aparılmış eksperimentlər seçilmiş zaman intervallarında əlamət əhəmiyyətlərinin oxşarlığını və verilənlərin paylanmasında kiçik dəyişiklikləri göstərərək monitorinq yanaşmasının tətbiq imkanlarını dəstəkləyir, lakin modelin gələcək dövrlərdə dayanıqlığının təsdiqi üçün onun yenidən öyrədilmədən müstəqil zaman əsaslı testlərdə və real iş şəraitində dəqiqlik, tamlıq və F1 göstəriciləri üzrə qiymətləndirilməsi tələb olunur.

DİSSERTASİYA İŞİNİN ƏSAS NƏTİCƏLƏRİ

Dissertasiya mövzusu üzrə aparılan tədqiqatlar prosesində qoyulmuş elmi məsələlər həll edilmiş və aşağıdakı əsas nəticələr alınmışdır:

1. Zərərli proqram təminatının aşkarlanması sahəsində mövcud beynəlxalq təcrübə sistemli şəkildə təhlil edilmiş, sahənin əsas inkişaf meyilləri, tətbiq olunan texnologiyalar, mövcud metodoloji məhdudiyyətlər və elmi-tədqiqat boşluqları müəyyənləşdirilmişdir. Müxtəlif elmi məlumat bazalarında indeksləşdirilən nəşrlər üzrə kəmiyyət və keyfiyyət göstəricilərinə əsaslanan bibliometrik və müqayisəli təhlillər aparılmışdır [3], [9], [12], [16], [18]. Tədqiqatın nəticələrindən kibertəhlükəsizlik təşkilatlarında müasir aşkarlama yanaşmalarının seçilməsi və inkişaf strategiyalarının müəyyənləşdirilməsi üçün istifadə oluna bilər.

2. Zərərli proqram təminatının dayanıqlı və effektiv aşkarlanmasını təmin edən üçmərhələli konseptual model işlənilmişdir. Təklif olunan model zərərli proqram nümunələrinin ilkin emalı, aşkarlama modelinin qurulması və real istismar mühitində tətbiqi mərhələlərini əhatə etməklə yanaşı, model qərarlarının izahediləbilənliyi, fəaliyyət göstəricilərinin monitorinqini və dayanıqlığının yoxlanılmasını da nəzərdə tutur [21]. Üçmərhələli konseptual modelin Təhlükəsizlik Əməliyyat Mərkəzlərində tətbiqi ilkin emal, aşkarlama, qərarların izahı və monitorinq proseslərinin avtomatlaşdırılmasına xidmət edə bilər.

3. Maşın öyrənməsi modellərinin qurulması zamanı istifadə olunan statik və dinamik əlamətlər fəzasından informativ əlamətlərin seçilməsi üçün əlamətlərin diskret və kəsilməz xarakterini nəzərə alan hibrid filtrasiya metodu işlənilmişdir. Təklif olunan metod əlamətlər fəzasının ölçüsünün azaldılmasına, hesablama mürəkkəbliyinin aşağı salınmasına və aşkarlama prosesinin səmərəliliyinin yüksəldilməsinə imkan vermişdir [23]. Hibrid əlamət seçimi metodu hesablama yükünün azaldılması və böyük həcmli verilənlərin sürətli emalı üçün tövsiyə olunur.

4. Windows və Android əməliyyat sistemləri mühitində zərərli proqram təminatının, eləcə də zərərli şəbəkə trafikinin aşkarlanması üçün statik və dinamik analiz nəticəsində əldə edilmiş müxtəlif tipli

əlamətlərdən istifadə edən klassifikasiya və klasterləşdirmə əsaslı hiperparametrlərinin optimallaşdırılmasına, informativ əlamətlərin seçilməsinə və verilənlərin xüsusiyyətlərinə uyğunlaşdırılmasına əsaslanan aşkarlama metodları işlənmişdir. Təklif olunan metodların binar və çoxsinifli klassifikasiyası və oxşar nümunələrin klasterləşdirilməsi məsələlərində effektivliyi eksperimental olaraq qiymətləndirilmişdir [1], [2], [6-8], [10], [13]. Klassifikasiya və klasterləşdirmə metodları fərdi kompüterlərdə, Android cihazlarında və şəbəkələrdə zərərli proqramların və təhlükəli trafikin vaxtında aşkarlanmasında istifadə edilə bilər.

5. Zərərli proqram təminatının təsvirlərə çevrilərək aşkarlanması üçün transfer öyrənmə əsaslı yüksək dəqiqlikli aşkarlama modeli işlənilmişdir. Eksperiment nəticəsində MobileNet modeli yüksək (98.9% F1-ölçü) və dayanıqlı performans göstərmişdir [4], [11], [20]. MobileNet əsaslı yanaşma resursları məhdud cihazlarda zərərli kodların təsvirlər vasitəsilə təhlili üçün tətbiq oluna bilər.

6. Təlim verilənlərində zərərli proqram nümunələrinin sayının artırılması və siniflərarası balanssızlığın azaldılması məqsədilə şərti generativ rəqib şəbəkəyə əsaslanan sintetik nümunələrin generasiyası metodu işlənilmişdir. Generasiya edilmiş nümunələrlə zənginləşdirilmiş verilənlər toplusu üzərində maşın öyrənməsi modelləri sınaqdan keçirilmiş və onların aşkarlama göstəricilərində yüksəlmə müşahidə olunmuşdur [5], [12], [15]. Şərti generativ rəqib şəbəkə ilə sintetik nümunələrin yaradılması təlim bazalarında nadir zərərli proqramların sayını artırmağa və sinif balanssızlığını azaltmağa imkan verə bilər.

7. Təklif olunan zərərli proqram aşkarlama modellərinin qərarvermə mexanizmlərinin izahediləbilənliyinin qiymətləndirilməsi məqsədilə üç müxtəlif verilənlər toplusu üzərində eksperimental tədqiqatlar aparılmışdır. SHAP, LIME və əlamətlərin əhəmiyyətliliyinin qiymətləndirilməsi metodlarından istifadə etməklə ayrı-ayrı əlamətlərin model proqnozlarına təsiri müəyyənləşdirilmiş və modellərin qərarları interpretasiya edilmiş, həmçinin, izahetmə sabitliyini ölçən riyazi qiymətləndirmə metodu təklif edilmişdir [17], [19], [22]. Bundan əlavə, modellərin real mühitdə tətbiqi zamanı verilənlərin paylanmasında və əlamətlərin strukturunda baş verən

yayınmaların müəyyən edilməsi üçün populyasiya stabilliyi indeksi — PSI metodu eksperimental olaraq qiymətləndirilmişdir [14]. SHAP və LIME model qərarlarının şərhi, PSI isə verilənlərin paylanmasındakı dəyişikliklərin izlənməsi üçün istifadə edilə bilər. Bu vasitələr analitiklərin qərarların səbəblərini anlamasına və modellərin yenilənməsi ehtiyacını qiymətləndirməsinə və zəruri mühafizə tədbirləri seçməsinə kömək edə bilər.

Dissertasiya işinin əsas nəticələri aşağıdakı elmi işlərdə dərc edilmişdir:

1. Bağırov, E.O. Zərərli proqramların n-qram analizi əsasında aşkarlanması // Tələbə və gənc tədqiqatçıların I beynəlxalq elmi konfransının materialları, – Bakı, – 6–16 aprel, – 2020, – s. 431–433.
2. Bağırov, E.O. Klasterizasiya üsullarının zərərli proqramların aşkarlanmasına tətbiqi // Tələbə və gənc tədqiqatçıların II beynəlxalq elmi konfransının materialları, – Bakı, – 13–28 aprel, – 2021, – s. 438–440.
3. Baghirov, E. Techniques of Malware Detection: Research Review // Proceedings of the 2021 IEEE 15th International Conference on Application of Information and Communication Technologies (AICT), – Baku: IEEE, – 13–15 October, – 2021, – p. 1–6. **(Scopus, WoS)**
4. Baghirov, E. Image-Based Malware Detection Method // Proceedings of the II International Conference on Information Security: Problems and Prospects, – Baku: Azerbaijan University, – 25 November, – 2022, – p. 28–31.
5. Bağırov, E.O. Zərərli proqramların generativ rəqib şəbəkələrlə generasiyası // "Rəqəmsal Transformasiya–2023" IV beynəlxalq elmi konfransının materialları, – Bakı, – 18–19 aprel, – 2023, – s. 378–380.
6. Baghirov, E. Machine Learning-based Android malware detection: a robust and accurate approach // Proceedings of the VII International Scientific Conference of Young Researchers, – Baku, – 28–29 April, – 2023, – p. 1–10.

7. Baghirov, E. Leveraging Machine Learning for Accurate Malware Traffic Detection // Proceedings of the International Conference on Intelligent Systems and New Applications (ICISNA'23), – Liverpool, – 28–30 April, – 2023, – p. 25–26.
8. Baghirov, E. Evaluating the Performance of Different Machine Learning Algorithms for Android Malware Detection // Proceedings of the 2023 5th International Conference on Problems of Cybernetics and Informatics (PCI), – Baku: IEEE, – 28–30 August, – 2023, – p. 1–4. **(Scopus, WoS)**
9. Baghirov, E.O. Analyzing the performance of behavioral-based malware detection approaches under real-world conditions // Optical-Electronic Devices and Devices in Pattern Recognition and Image Processing Systems: Proceedings of the XVII International Scientific and Technical Conference, – Kursk: South-West State University, – 12–15 September, – 2023, – p. 15–17. (**РИНЦ)**
10. Baghirov, E. Comprehensive Framework for Malware Detection: Leveraging Ensemble Methods, Feature Selection and Hyperparameter Optimization // Proceedings of the 2023 IEEE 17th International Conference on Application of Information and Communication Technologies (AICT), – Baku: IEEE, – 18–20 October, – 2023, – p. 1–5. **(Scopus, WoS)**
11. Baghirov, E. Disassembled files to grayscale images: Advancing malware detection with diverse features // Doktorantların və gənc tədqiqatçıların XXVI respublika elmi konfransının materialları, – Bakı, – 10–11 noyabr, – 2023. Cild 1, – p. 130–134.
12. Baghirov, E.O. Critical analysis and review of current research on GANs for malware detection // Proceedings of the 12th National Supercomputer Forum, – Pereslavl-Zalessky, – 28 November–1 December, – 2023, – p. 1–10.
13. Baghirov, E. Malware detection based on opcode frequency // Problems of Information Technology, – 2023. 14(1), – p. 3–7.
14. Baghirov, E. A comprehensive framework for real-time malware detection and monitoring in production // International Journal on Information Technologies and Security, – 2024, – 16(4), – p. 85-94. **(WoS)**

15. Baghirov, E. Building robust malware detection through Conditional Generative Adversarial Network-based data augmentation // Program Systems: Theory and Applications, – 2024. 15(4), – p. 97–110. (**WoS**, **РИНЦ)**
16. Baghirov, E., Afzal, M. Unveiling Malicious Patterns: Autoencoder-Based Malware Detection // System Informatics, – 2024. 24, – p. 113–122. **(РИНЦ)**
17. Baghirov, E. Advanced machine learning and interpretability for Windows malware detection // Journal of Modern Technology and Engineering, – 2024. 9(3), – p. 165–177.
18. Imamverdiyev, Y.N., Baghirov, E.O. Evasion Techniques in Malware Detection: Challenges and Countermeasures // Problems of Information Technology, – 2024. 15(2), – p. 9–15.
19. Imamverdiyev, Y., Baghirov, E., Ikechukwu, J. Detecting obfuscated malware infections on Windows using ensemble learning techniques // Informatics and Automation, – 2025. 24(1), – p. 99–124. **(Scopus, WoS)**
20. Imamverdiyev, Y., Baghirov, E., Ikechukwu, J.C. Image-based Deep Learning method for effective malware detection // Problems of Artificial Intelligence, – 2025. № 1(36), – p. 106–122. **(РИНЦ)**
21. Bağırov, E.O. Zərərli proqramların aşkarlanması sisteminin ümumi strukturu // "Mütərəqqi texnologiyalar və innovasiyalar" X respublika elmi-texniki konfransının materialları, – Bakı, – 1–2 may, – 2025, – s. 310–313.
22. Baghirov, E. A comprehensive investigation into robust malware detection with explainable AI // Cyber Security and Applications, – 2025. 3, – Article 100072. **(Scopus – Q1)**
23. Bağırov, E.O. Zərərli proqramların aşkarlanması modelləri üçün əlamətlərin seçilməsi üsulu // "Heydər Əliyevin dövlət təhlükəsizliyi siyasəti: tarixi nailiyyətlər və müasir çağırışlar" respublika elmi-praktik konfransının materialları, – Bakı, – 6 may, – 2025, – s. 294–299.

Həmmüəlliflərlə dərc olunmuş işlərdə iddiaçının şəxsi rolu:
[16]–Verilənlər toplusunun hazırlanması; metodologiya; vizuallaşdırma; nəticələrin interpretasiyası.
[18]– İcmal; məqalənin mətninin yazılması.
[19]–Metodologiya; eksperimentlər; vizuallaşdırma; nəticələrin interpretasiyası; məqalənin mətninin yazılması.
[20]–Metodologiya; eksperimentlər; vizuallaşdırma; nəticələrin interpretasiyası; məqalənin mətninin yazılması.



Dissertasiyanın müdafiəsi 27 noyabr 2026-cı il tarixində saat 16^{00}-da Azərbaycan Respublikası Elm və Təhsil Nazirliyinin İnformasiya Texnologiyaları İnstitutunun nəzdindəki FD 1.35 Dissertasiya Şurasının iclasında keçiriləcək.

Ünvan: Az1141, Bakı şəhəri, B.Vahabzadə küçəsi, 9A.

Dissertasiya ilə Azərbaycan Respublikası Elm və Təhsil Nazirliyinin İnformasiya Texnologiyaları İnstitutunun kitabxanasında tanış olmaq mümkündür.

Dissertasiya və avtoreferatın elektron versiyaları Azərbaycan Respublikası Elm və Təhsil Nazirliyinin İnformasiya Texnologiyaları İnstitutunun rəsmi internet saytında (ict.az) yerləşdirilmişdir.

Avtoreferat 26 oktyabr 2026-cı il tarixində zəruri ünvanlara göndərilmişdir.

Çapa imzalanıb: 22.10.2026

Kağızın formatı: A5

Həcm: 39900 işarə

Tiraj: 100