AZƏRBAYCAN RESPUBLİKASI ELM VƏ TƏHSİL NAZİRLİYİ

AZƏRBAYCAN DÖVLƏT İQTİSAD UNİVERSİTETİ

UNEC BİZNES MƏKTƏBİ

Biznes Məktəbinin direktoru

i.e.n., dos. Hacıyev Nazim Özbəy

"_____" (imza)

"___"_____2023-cü il

"KİBERTƏHLÜKƏSİZLİKDƏ SÜNİ İNTELLEKTƏ YÖNƏLƏN HÜCUMLARIN TƏHLİLİ"

MAGİSTR DİSSERTASİYA İŞİ

İxtisasın şifri və adı: 060409 Biznesin idarə edilməsi

İxtisaslaşma: Kibertəhlükəsizlik

Qrup: A19-21

Magistrant	Elmi rəhbər
Rəcəbli Sevinc Şirazi	t.e.f.d. Əbdiyeva-Əliyeva Günay Əlişan
_____imza	_____imza
Proqram rəhbəri	**Kafedra müdiri**
i.f.d. Qəzənfərli Mirvari Xəqani	dos. Məmmədova Sevər Mömin
_____imza	_____imza

BAKI - 2023

Təşəkkürnamə

Təşəkkürnamə vasitəsilə dissertasiya işi ilə bağlı tədqiqatların aparılması və həmin tədqiqatlar əsasında nəzəri-praktiki olmaqla biliklərin əldə edilməsi yönündə dəstək göstərmiş olan, t.e.f.d., Günay Əbdiyeva-Əliyeva, Bayraməli Nəzərov və Ramil Hüseynova təşəkkürümü bildirir, eyni zamanda praktiki olaraq tədqiqat işinin araşdırılması üçün avadanlıq təminatının və araşdırma mühitinin təşkilində dəstək göstərən “Cybernetics” MMC rəhbərliyinə xüsusi təşəkkürlərimi bildirirəm.

KİBERTƏHLÜKƏSİZLİKDƏ SÜNİ İNTELLEKTƏ YÖNƏLMİŞ HÜCUMLARIN TƏHLİLİ

Xülasə

Dissertasiya işinin aktuallığı: 2023-cü il üçün qeydə alınan ən aktual süni intellekt əsaslı təhdidlərin vaxtında aşkarlanması və aradan qaldırılması aktualdır.

Dissertasiya işinin məqsədi: Süni intellektə yönəlmiş təhdid, zəiflik və istismarların qarşısının alınması üçün erkən xəbərdarlıq sistemlərinin və qarşısının qabaqcıl texnologiya ilə vaxtında operativ və effektiv alınması əsas məqsəddir.

Dissertasiya işində istifadə edilmiş biznes tədqiqat metodları: Süni intellektin tətbiqi ilə IDS və IPS - müdaxilənin aşkarlanması və qarşısının alınması üsulları ilə "CYBERNETİCS" şirkətinin server və mail domenləri üzərində tədqiqat aparılmışdır.

Dissertasiya işinin informasiya bazası: Müxtəlif internet saytlarında dərc olunan bloq yazı nümunələrindən, türk və ingilis dilində çap edilmiş kitab və elmi məqalələrdən istifadə edilmişdir.

Dissertasiya işinin məhdudiyyətləri: Artan təhdidlərin qarşısının alınması üçün sahə üzrə mütəxəssislərin mütəmadi ixtisasartırma kurslarına cəlb olunması, tədqiqat işlərinə maliyyə vəsaitinin ayrılması təşkilatların prioritetlərindən olmalıdır.

Dissertasiya işinin praktik nəticələri: Yeni nəsil təhlükəsizlik divarlarında maşın öyrənmə alqoritmlərindən istifadə edərək baş verə biləcək təhdidlərə qarşı həssaslığını öyrənməkdən və qarşısını alma tədbirlərindən istifadə etməkdən ibarətdir.

Açar sözlər: **Süni intellekt, APT, kibertəhlükəsizlik, XGBoost, DDoS, DoS, AI, Artiffical Inteligence.**

MÜNDƏRİCAT

Giriş

Mövzunun aktuallığı: Bildiyimiz kimi informasiya texnologiyaları və informasiya sistemlərinin təhlükəsizliyi gündən-günə daha da inkişaf etməktədir. İstər Azərbaycanda, istərsə də digər ölkələrdə texnoloji inkişaf ildən-ilə yeniliklər gətirməkdə davam edir. Rəqəmsal transformasiya dediyimiz bu dövr ərzində bir çox təkamül mərhələləri getmiş və elmi yeniliklər insanlara istifadə üçün təqdim edilmiş və ya onların gələcəkdə mümkünlük dərəcəsi bildirilmişdir.

Dissertasiya işinin məqsədi: Tədqiqatın aparılmasında əsas məqsəd qabaqcıl süni intellekt texnologiyalarının inkişafının gətirə biləcəyi boşluq və təhdidlərin öncədən praqnozlaşdırılması və onlara qarşı mübarizə mexanizmlərini yaratmadan öncə süni intellektin daha dərindən mənimsənilməsidir. Maşın öyrənmə və XGBoost alqoritmləriylə, eyni zamanda bir çox SİEM vasitələri, müdaxilənin aşkarlanması sistemləri ilə qabaqcıl müdaxilə testləri həyata keçirmək üçün bir sıra analizlərin aparılmasına əsaslanır.

Tədqiqatın predmeti: Günümüzdə təhsil sahəsində, səhiyyədə xüsusilə də neyrocərrahiyədə, maliyyədə, aviasiyada və s. buna bənzər digər sahələrdə tətbiqi nəzərdə tutulan süni intellekt gələcək zamanda qloballaşmış müasir texnoloji trend olacağı proqnozlaşdırılır. Amerika Birləşmiş Ştatları, Çin, Cənubi Koreya, bir neçə Avropa ölkəsi bu gün üçün süni intellektin mərkəzi hesab olunur. Adı sadalanan ölkələrdə təhsil, səhiyyə, ordu, maliyyə, neft və qaz sənayesində, metallurgiya sektorunda, müxtəlif ticarət sahəsi və nəqliyyat sektorunda tətbiq olunan süni intellekt olduqca perspektivli hesab edilir.

Tədqiqatın obyekti: Süni intellekt gələcəyin sahəsi olmaqla, yaxın on illik ərzində getdikcə aktuallaşması gözlənilən sahədir. Bir çox informasiya texnologiyaları və informasiya təhlükəsizliyi ilə məşğul olan şirkətlərlə savayı, bir

çox qeyri-ixtisas sahələrində də qabaqcıl texnologiya şəklində istifadəsi nəzərdə tutulmaqdadır.

Elmi yenilik: Texnoloji olaraq, bu inqilabın mərkəzi hissəsində yerləşməkdə olan süni intellekt nəinki, informasiya texnologiyaları sahəsində, eləcə də avtomobil, səhiyyə, bank sahələrində də sürətli inkişaf edəcəyi gözlənilir.

İşin praktiki əhəmiyyəti: Tədqiqat işi zamanı süni intellekt və maşın öyrənmə, eyni zamanda IDS və IPS sistemlərinin istifadəsi ilə bağlı təklif və tövsiyələri, aktiv olan kiber müdafiə metodları ilə birlikdə həyata keçirilməsi, kiber mühitlərdə, həmçinin müəssisələrdə təhlükəsizliyin planlaşdırılma və tətbiqinə, o cümlədən təkmilləşdirilməsi üçün böyük yardım ola bilər.

Tədqiqatın elmi və nəzəri əsasları: Tədqiqat zamanı mövzu ilə əlaqəli nəzəri əsaslar hazırlanarkən xarici ölkələrin, o cümlədən Türkiyənin mütəxəssisləri tərəfindən tədqiq və tətbiq olunmuş resurslardan və başqa mənbələrdən istifadə edilərək dissertasiya işinin mövzusu barədə məlumatlar dərindən öyrənilərək fərqli tətbiq metodları ilə tanış olunmuşdur.

İşin struktur və həcmi: Tədqiqat işində giriş, üç fəsil, nəticə və təkliflər, ixtisar olunmuş sözlərin açıqlaması və dissertasiya işinin tədqiqi zamanı istifadə edilmiş ədəbiyyat siyahısı əks olunmuşdur.

Birinci fəsildə, üç alt başlıqdan təşkil edilmiş,kompüter sistemlərinə yönəlik hücum növləri, qarşılaşıla biləcək sosial mühəndislik taktikaları, deepfake və onları reallaşdıra biləcək pisniyyətli hücumçular, statistik göstəricilər, qarşı-qarşıya qala bilinəcək təhdidlər və həmin təhdidlərin qarşısını alması üçün reallaşdırıla bilinəcək metodlar haqqında məlumatlar öz əksini tapmışdır.

İkinci fəsildə süni intellektə daxil olan ekspert sistemlərdən söz açmaqla, hər hansısa bir sahədə bir mütəxəssisin yerinə yetirə biləcəyi qərar qəbuetmə prosesini təqlid etmə barədə danışılmış, bununla yanaşı ağıllı agentlərin iş prinsipindən bəhs edilmişdir. Sonda isə genetik alqoritmlər və süni intellektdə mahiyyəti əks olunmuşdur.

Üçüncü fəsilə gəldikdə isə süni intellekt hücumları, onların yarada biləcəyi fəsadların nümunələri, maşın öyrənmə alqoritmləri ilə performans analizinin təhlili, son olaraq isə zorla müdaxilənin aşkarlanması sistemləri və hücumları süni intellektə tanıtmaqla onların qarşısının alınma üsullarından bəhs edilmişdir.

FƏSİL 1. KOMPÜTER SİSTEMLƏRİNƏ YÖNƏLMİŞ HÜCUMLAR VƏ ONLARDAN MÜHAFİZƏ MEXANİZMLƏRİNİN İŞLƏNMƏSİ

1.1 Müasir dövrdə kompüter sistemlərinə yönəlmiş hücumların təhlili

Son dövrlərdə kompüter sistemlərinə edilən hücumların təhlili getdikcə daha çox əhəmiyyət kəsb etməyə başlayıb. Bu hücumlar bir çox sənaye sahələrində - şirkətlərdə, qurumlar və hökumətlər üçün məlumatların təhlükəsizliyi və məxfiliyi ilə bağlı narahatlıqları artırmaqdadır.

Hücum təhlili hücumun uğurlu olub-olmadığını, niyə uğurlu olduğunu və haradan gəldiyini müəyyən etmək məqsədi daşıyır. Bu analiz hücumların qarşısını almaq, aşkar etmək və dayandırmaq üçün istifadə edilə bilər.

Müasir dövrdə kompüter sistemlərinə hücumlar çox vaxt hakerlər və ya kibercinayətkarlar tərəfindən həyata keçirilir. Bu hücumların əksəriyyəti kompüter şəbəkələrinə və ya sistemlərinə giriş əldə etmək üçün istifadə edilən zəiflikləri hədəf alır. Məsələn, istismar, zərərli proqram və ya sosial mühəndislik kimi üsullardan istifadə edilə bilər.

Kompüter sistemlərinə hücumlar bu gün artan təhlükədir. Bu hücumlar kibercinayətkarların və hakerlərin bir çox sənaye, biznes və hökumətlərdə məlumat təhlükəsizliyi və məxfilik mövzusunda narahat olmasına səbəb olur. Buna görə də kompüter sistemlərinə edilən hücumları təhlil etmək son dərəcə vacibdir. Yazımın davamında kompüter sistemlərinə hücumların müasir təhlili müzakirə olunacaq.

Kompüter sisteminə yönəlik hücum kibercinayətkarlar və hakerlər kompüter şəbəkələrinə və ya sistemlərinə giriş əldə etmək üçün istifadə edilən zəiflikləri hədəfə aldıqda baş verdiyindən, bu hücumların müxtəlif üsulları meydana gəlir.

Zəifliklər hakerlər tərəfindən kompüter sistemlərindəki boşluqlardan istifadə etmək üçün araşdırılır. Bu boşluqlar kibercinayətkarlara və ya hakerlərə kompüter sistemlərinə nəzarəti ələ keçirməyə və ya məlumatları oğurlamağa imkan verir.

Zərərli proqram kompüter sistemini yoluxduran zərərli proqramdır. Bu proqramlar kompüter sistemlərinə zərər vermək üçün nəzərdə tutulub. Sosial mühəndislik kompüter sistemlərinə daxil olmaqla insanların təhlükəsizlik boşluqlarından istifadə etmək üçün istifadə olunur. Məsələn, kibercinayətkarlar kiməsə saxta e-poçt göndərməklə onu aldada bilərlər.

Hücum və təhdidlərin təhlilinin məqsədi hücumun uğurlu olub-olmadığını, niyə uğurlu olduğunu və hücumun haradan gəldiyini müəyyən etməkdir. Bu analiz hücumların qarşısını almaq, aşkar etmək və bloklamaq üçün istifadə edilə bilər. Bu cür analiz bir çox müxtəlif texnika və vasitələrdən istifadə etməklə edilə bilər. Məsələn, log girişləri hücumları izləmək üçün istifadə edilə bilər. Bu jurnallar kompüter sistemində baş verən hadisələri qeyd edir və bu hadisələri təhlil etməyə imkan verir. Hücumları aşkar etmək üçün istifadə edilən digər üsul şəbəkə monitorinq alətləridir. Bu alətlər trafikə və kompüter şəbəkələrinə potensial hücumlara nəzarət edir. Bu alətlər həmçinin şəbəkədəki cihazların vəziyyətini və yeniləmələrini izləyə bilər. Bundan əlavə, kompüter sistemlərinə hücumları aşkar etmək üçün antivirus proqramları və firewall-lardan istifadə edilə bilər.

Hücumun mənbəyini müəyyən etmək üçün istifadə edilən digər üsul hücumun həyata keçirildiyi IP ünvanını aşkar etmək və təhlil etməkdir. Bu təhlil hücumun mənbəyini müəyyən etmək üçün istifadə olunur və bu məlumat hücumların qarşısını almağa və gələcəkdə oxşar hücumları aşkar etməyə kömək edir. Hücumların qarşısını almaq üçün istifadə edilən digər üsul hücuma məruz qalan kompüter sisteminin və ya şəbəkənin təhlükəsizlik tədbirlərini söndürmək və ya artırmaqdır. Bu, gələcək hücumların qarşısını almaq və sistemi təhlükəsiz saxlamaq üçün lazımdır. Kompüter sistemlərinə hücumlar bu gün artmaqda olan təhlükədir. Buna görə də hücumların təhlili son dərəcə vacibdir. Hücum analitikası hücumları aşkar etməyə, tapmağa və qarşısını almağa kömək edir. Şəbəkə təhlükəsizliyində bir çox fərqli texnika və vasitələrdən istifadə edərək hücum təhlili son dərəcə vacibdir. Təşkilatlar və şəxslər

kompüter sistemlərinə edilən hücumların qarşısını almaq üçün təhlükəsizlik tədbirləri görməli və hücumları təhlil etmək üçün mütəxəssislərlə əməkdaşlıq etməlidirlər. Pisniyyətli şəxslər zərər vermək məqsədi ilə yaxud, kompüter sistemlərinə hücum etmək və ya komputer şəbəkələrinə daxil olaraq zərər vurmağa cəhd edirlər.

Müxtəlif üsullarla kompüter sistemlərinə qarşı hücumlar həyata keçirmək mümkündür. Bir neçə növ hücum var, adətən bu belə görünür:

-Kompüter sistemində zərərli hərəkətlər edən viruslar onu yoluxduran proqramlardır.

-Digər kompüterlərə yayılan qurdlar kompüter şəbəkələrini yoluxduran proqramlardır.

-İstifadəçinin xəbəri və ya razılığı olmadan quraşdırılmış troyan proqram təminatı kompüter sistemlərində gizlədilə bilər.

-Şəxsi məlumatları oğurlamağa çalışan zərərli e-poçtlardan və ya vebsaytlardan istifadə edir ki, bu da fişinq kimi tanınır.

-Kompüter şəbəkələrini hədəf alan Xidmətdən imtina (Denial of Service) – DoS hücumu sistemi sıx trafiklə doldurmaqla həyata keçirilir və nəticədə onun normal funksiyaları dayandırılır.

Optimal təhlükəsizlik tədbirləri üçün etibarlı təhlükəsizlik proqramlarından istifadə etmək, cihazlarını səylə yeniləmək və güclü parollardan istifadə etmək tövsiyə olunur. Əlavə olaraq, həm e-poçtların, həm də veb saytların qanuniliyini yoxlamaqla özünüzü qorumaq, naməlum mənbələrdən faylları açmaqdan çəkinməklə yanaşı, istifadəçi təhlükəsizliyi üçün çox vacibdir.

Hücumların təsnifatı bəzi təhlil tələb edən çətin bir mövzu ola bilər. Hücumlar müxtəlif formalarda ola bilər, ona görə də qruplaşdırma metodu onları anlamağa kömək edə bilər.

Bu təsnifat çərçivəsində kiber və fiziki hücumları ayırd etmək faydalı ola bilər. Bundan əlavə, hücumlar məqsədlərinə görə təsnif edilə bilər, məsələn, casusluq və ya təxribat buna aiddir. Hücumların təsnifləşdirilməsinin bəzi digər üsullarına maliyyə

sistemləri və ya məlumat mərkəzləri kimi hansı növ hədəfin hücum edildiyi daxildir. Hücumları bu şəkildə təşkil etməklə, nümunələri müəyyən etmək və gələcək təhlükələri təxmin etmək daha asan ola bilər. Bununla belə, təsnifatın özünü müəyyən etmək bəzən çətin ola bilər, çünki bəzi hücumlar bir neçə məqsədə malik ola bilər və ya bir neçə domeni hədəfləyə bilər.

Ümumiyyətlə, kompüter sistemi hücumlarının bir neçə kateqoriyasına aid edilir.

Proqram təminatına edilən hücumlar əsl problem ola bilər. Bu tip hücumlar kompüterinizə ciddi təsir göstərə bilər və heç bir xəbərdarlıq edilmədən baş verə bilər. Çox gec olana qədər hücumun baş verdiyini belə dərk etməyə bilərsiniz. Yadda saxlamaq lazım olan ən vacib şey odur ki, proqram təminatı zəiflikləri təcavüzkarlar tərəfindən istifadə edilə bilər. Onlar zəif yazılmış koddan istifadə edə və ya proqram təminatınızda təhlükəsizlik zəifliklərindən istifadə yollarını tapa bilərlər. Buna görə də bütün proqram təminatınızın yeni olduğundan və hər şeyin ən son versiyalarından istifadə etdiyinizə əmin olmaq çox vacibdir. Əks təqdirdə, özünüzü çox qısa müddətdə çoxlu zərər verə biləcək hücumlara qarşı həssas edə bilərsiniz.

Proqram təminatı kompüter sistemlərini yoluxdurur və hücumlar zamanı zərərli prosesləri həyata keçirir. İstifadə olunan müxtəlif zərərli proqram növlərinə zərərli proqramlar, viruslar, qurdlar və troyanlar daxildir.

Kiberhücumun bir növü olan şəbəkə əsaslı hücumlar bağlı sistemlərə ciddi ziyan vura bilər. Bu hücumlar kompüter sistemləri və ya internet əlaqələri kimi şəbəkələri pozmaq məqsədi daşıyır və zərərli şəxslər və ya qruplar tərəfindən həyata keçirilə bilər. Kibercinayətkarlar çox vaxt həssas məlumatlara daxil olmaq və ya şəbəkə trafikini manipulyasiya etmək üçün zərərli proqram, fişinq və ya xidmətdən imtina (DoS) hücumları kimi müxtəlif üsullardan istifadə edirlər. Şəbəkəyə əsaslanan hücumlardan qorunmaq üçün təşkilatlar firewall, şifrələmə və müdaxilənin aşkarlanması sistemləri kimi möhkəm təhlükəsizlik tədbirlərini həyata keçirməlidir.

Ayıq qalmaq və ən son təhlükələr barədə məlumatlı olmaq şəbəkənizi təhlükəsiz saxlamaq üçün vacibdir.

Şəbəkə trafiki üzərində baş verə biləcək bir neçə növ hücum var. Bu cür hücumlar Paylanmış Xidmətdən imtina (DDoS), Meet-in-the-Middle (MitM) hücumları və Xidmətdən imtina (DoS) hücumlarından ibarətdir.

Təcavüzkarlar e-poçt hesablarına sızmaq və saxta mesajlar göndərmək üçün müxtəlif üsullardan istifadə etdikləri üçün mesajlaşma və e-poçt təhdidləri artır. Bu hücumlar son dərəcə mürəkkəb və aşkarlanması çətin ola bilər, potensial olaraq həssas məlumatlara və ya şəxsi məlumatlara təhlükə yarada bilər. Hesablarınızı və məlumatlarınızı bu hücumlardan qorumaq üçün ayıq olmaq və lazımi tədbirləri görmək vacibdir. Spama və ya e-poçta etibar etməyib və həmişə güclü parollara həmçinin, iki faktorlu autentifikasiyaya sahib olmaq lazımdır. Ayıq və fəal olmaqla, bu zərərli hücumların bizə və ya təşkilatımıza təsir etməsinin qarşısını ala bilərik.

Bu cür hücumlar istənməyən mesajlar və ya spam e-poçtlar vasitəsilə baş verə bilər. Onların tərkibində zərərli fayllar və ya keçidlər ola bilər.

Özümüzü arzuolunmaz hücumlardan qorumaq bir sıra ehtiyatlı tədbirlərlə əldə edilə bilər, o cümlədən şübhəli mənbələrdən olan fayllara girişdən qaçınmaq, yenilənmiş təhlükəsizlik proqramından istifadə etmək, təhlükəsiz və uzun parolların tətbiqi də bunlara daxildir.

1.2 Mənbə kod istismarının analizi (Code exploit)

Zərərli kod və ya terminoloji adı ilə desək, “code exploit” icazəsiz giriş əldə etmək üçün kompüter sistemlərindəki səhvlərdən istifadəni nəzərdə tutan və kod istismarından istifadə edərək təcavüzkarlar tərəfindən icra edilə bilən bir koddur.

Çox vaxt kompüter sistemi təcavüzkarlar tərəfindən kod istismarından istifadə edərək pozula bilər. İstismarın hədəfi təcavüzkara hədəf kompüter sisteminə giriş imkanı verən kompüter sistemindəki zəiflik və ya zəiflikdir.

Kod zəifliklər kompüter sistemi proqramını yeniləmək, güclü parollardan istifadə etmək və yaxşı təhlükəsizlik proqramından istifadə etməklə azalda bilər. Bundan əlavə, naməlum mənbələrdən fayl və keçidlərin açılmaması kod hücumlarının qarşısının alınmasında mühüm addımdır.[10]

Kod istismarı ümumi hücum üsulu olsa da, qorunma üsulları var. Proqram tərtibatçıları proqram təminatında zəifliklər minimuma endirmək üçün təhlükəsizlik testi və kod auditi keçirə bilər. Onlar həmçinin bu zəifliklər yeniləmələrlə aradan qaldırmalıdırlar.

Nəticə etibariylə, istismar hücumları kompüter sistemlərində ciddi boşluqlara səbəb ola bilər və bu zəifliklərin aşkarlanması, istismarı və qorunması çox vacibdir. Proqram tərtibatçıları təhlükəsizlik pozuntularını minimuma endirmək üçün lazımi tədbirlər görməlidirlər və istifadəçilər kompüterlərini qorumaq üçün təhlükəsizlik proqramlarını yeniləməlidirlər.

Gizli dinləmə prosesinin aşkarlanması. Dinləmə hücumu təcavüzkarın icazəsiz bir şəxsin və ya təşkilatın şəxsi söhbətlərini, mesajlarını və ya digər ünsiyyətlərini dinləməsidir. Bu cür hücumlar şəxsi məxfiliyi poza və ciddi təhlükəsizlik riskləri yarada bilər.

Dinləmə hücumları adətən dinləmə qurğularının quraşdırılması ilə həyata keçirilir. Bu dinləmə cihazları telefon xətlərinə, internet bağlantılarına və ya simsiz şəbəkələrə yerləşdirilə və hədəfin ünsiyyətinə qulaq asa bilər. Bundan əlavə, təcavüzkarlar hədəfin kompüterinə zərərli proqram quraşdırmaqla və ya onun Wi-Fi şəbəkəsinə daxil olmaqla hədəfin şəxsi kommunikasiyalarına icazəsiz giriş əldə edə bilərlər.[11]

Şəkil 1. Gizli dinləmə prosesinin təsviri

Mənbə. Müəllif tərəfindən tərtib edildi

Telefon dinləməsi fərdlər və ya təşkilatlar tərəfindən şəxsi söhbətlərin və ya mesajların icazəsiz dinlənməsi və ya izlənilməsidir. Bu söz hərfi mənada evin divarlarından kənarda səsə qulaq asmaq mənasını verir.

Dinləməni həm də şəbəkə təhlükəsizliyinə böyük təhlükə yaradan gizli dinləmə hücumu adlandırmaq olar. Məsələn, təcavüzkar hədəfin telefon zənglərini, mətn mesajlarını və ya veb trafikini onların razılığı olmadan dinləyə, şəxsi məlumatlarını və həssas məlumatlarını əldə edə bilər. Bu tip hücumlar kibercinayətkarlar tərəfindən tez-tez istifadə olunur və qurumlar, hökumətlər və fərdlər üçün ciddi təhlükəsizlik riski yaradır.

Dinləmələrin qarşısını almaq üçün insanlar güclü parollardan istifadə etməli, cihazlarını yeniləməli və təhlükəsizlik proqramlarından istifadə etməlidirlər. Həmçinin, açıq Wi-Fi şəbəkələrindən istifadə etməməli və yalnız etibarlı şəbəkələrə qoşulmalıdırlar. Həssas məlumatları ötürərkən şifrələmə də istifadə edilməlidir.

Arxa qapı (Backdoor)

Arxa qapı hücumu, təcavüzkarın icazəsiz giriş əldə etmək və nəzarəti əldə etmək üçün kompüter sistemindəki və ya şəbəkədəki zəiflikdən istifadə etdiyi hücumdur. Bu cür hücumlar adətən proqram təminatının boşluqları vasitəsilə və ya zərərli şəxs kompüter sisteminə daxil olduqda həyata keçirilir. Təcavüzkarlar bu zəiflikdən arxa qapılar kimi tanınan gizli giriş nöqtələri yaratmaq üçün istifadə edirlər. Bu giriş nöqtəsi ilə təcavüzkar istənilən vaxt hədəf kompüter sisteminə və ya şəbəkəyə icazəsiz giriş əldə edə və nəzarəti ələ keçirə bilər.

Arxa qapı və hücumu hücumçuya məlumatları oğurlamaq, dəyişdirmək və ya məhv etmək üçün hədəflənmiş kompüter sistemini və ya şəbəkəsini davamlı olaraq izləməyə imkan verir. Hücumçular həmçinin hədəflənmiş sistemlərə və ya şəbəkələrə zərərli proqram quraşdırmaq üçün arxa qapılardan istifadə edə bilərlər.

Şəkil 2. Arxa qapı (Backdoor) hücumunun təmsili

Mənbə: Müəllif tərəfindən tərtib edildi

Arxa qapı hücumlarının qarşısını almaq üçün istifadəçilər təhlükəsizlik zəifliklərini ciddi şəkildə izləməli və yeniləmələri vaxtında tətbiq etməlidirlər. Güclü parollardan istifadə etmək, etibarlı antivirus və təhlükəsizlik proqramlarından istifadə etmək,

yalnız etibarlı mənbələrdən proqram yükləmək və naməlum faylları açmamaq da vacibdir.

1.3 Sosial mühəndislik hücumları və müdafiə mexanizmlərinin işlənməsi

Sosial mühəndislik insanların davranışlarını manipulyasiya edərək və ya aldadaraq şəxsi məlumat və ya həssas məlumatları əldə etmək strategiyasıdır. Tez-tez hakerlər və zərərli təhlükə aktyorları tərəfindən istifadə edilən social mühəndislik insanların diqqətini yayındıraraq təhlükəsizlik tədbirlərindən yayınmaq üçün müxtəlif üsullardan istifadə edir.[12]

Sosial mühəndislik kiberhücumun bir hissəsi kimi istifadə edilə bilər. Məsələn, təcavüzkar hədəfin parolunu və ya digər məlumatları əldə etmək üçün təşkilatın və ya fərdin kompüter sistemlərinə daxil olmaq üçün sosial mühəndislik üsullarından istifadə edə bilər. Sosial mühəndislik fırıldaqçılıq və ya fişinq məqsədləri üçün də istifadə edilə bilər.

Şəkil 3. Sosial mühəndislik metodları.

Mənbə: Müəllif tərəfindən tərtib edildi

Sosial mühəndislik metodologiyalarına aşağıda sadalanan üsulları misal kimi göstərə bilərik:

Fişinq e-poçtları: Saxta ssenari əsasında hazırlanmış e-poçtlardan istifadə edərək, təcavüzkarlar insanların şəxsi məlumatlarını və ya parollarını oğurlamaq üçün sosial mühəndislik taktikalarından istifadə edə bilərlər.

Telefon fırıldaqçılığı: Təcavüzkarlar yalan hekayələr söyləyərək telefonda şəxsi və ya maliyyə məlumatlarını oğurlamaq üçün səlahiyyətli qurumlar və ya şəxslər kimi davrana bilərlər.

Sosial media fırıldaqçılığı: Sosial media platformalarından istifadə edərək, təcavüzkarlar şəxsi məlumat və ya pul oğurlamaq üçün saxta profillər və ya yalan hekayələr istifadə edərək insanların etibarını qazanmağa cəhd edə bilərlər.

Kripto fırıldaqçılıq: Təcavüzkarlar saxta vebsaytlardan və ya yalan vədlərdən istifadə edərək kriptovalyutaya investisiya etmək istəyən insanları hədəf alaraq kriptovalyuta oğurlamağa cəhd edə bilərlər.

Çiyin arxası izləmə: İstifadəçini ekrarına və klaviaturasına istifadə zamanı gizlincə baxaraq məxfi məlumatların ələ keçirilməsi

Zibil eşələmə: Haker özünə gərəkli məlumatı atılmış tullantıların arasından da əldə edə bilər.

Vişinq: Zəng zamanı istifadəçinin aldadılması ilə məlumatların oğurlanmasını həyata keçirir.

Fiziki yemləmə: Hədəfin marağını oyadaraq onun yemə düşməsini təmin etmir.

Başqa bir şey üçün verilən və ya alınan bir şey: Xidmət və ya fərqli faydalar qarşılığında sistemlərə giriş səlahiyyətlərinin istənilməsini həyata keçirir.

Bəhanə: Güvən qazanaraq zaman-zaman müxtəlif bəhanələrlə fərqli məlumatların istənilməsiylə məlumat oğurluğunun həyata keçirilməsini reallaşdırır.

Quyruq atma (Tailgating): Biometrik giriş kartlarının klonlanması və ya fərqli metodlardan istifadə edərək səlahiyyətsiz bölmələrə fiziki giriş imkanının yaradılmasıdır.

Su hövzəsi (Waterholing): Saytlarda gəzərkən mənfi xarakterli proqramların injeksiya olunmüş saytlar ilə kompyuterdən məlumatların oğurlanması prosesidir.

Əks sosial mühəndislik: Bu metodda hədəf özü hakere köməklik və ya başqa məqsədə nail olmaq üçün müraciət edir və haker göstərdiyi xidmət əsnasında məlumat oğurlayır.

Şəkil 4. Sənaye sahələrində e-poçtlara ən çox yönəlik fişinq hücum növlərinin faiz göstəricisi.

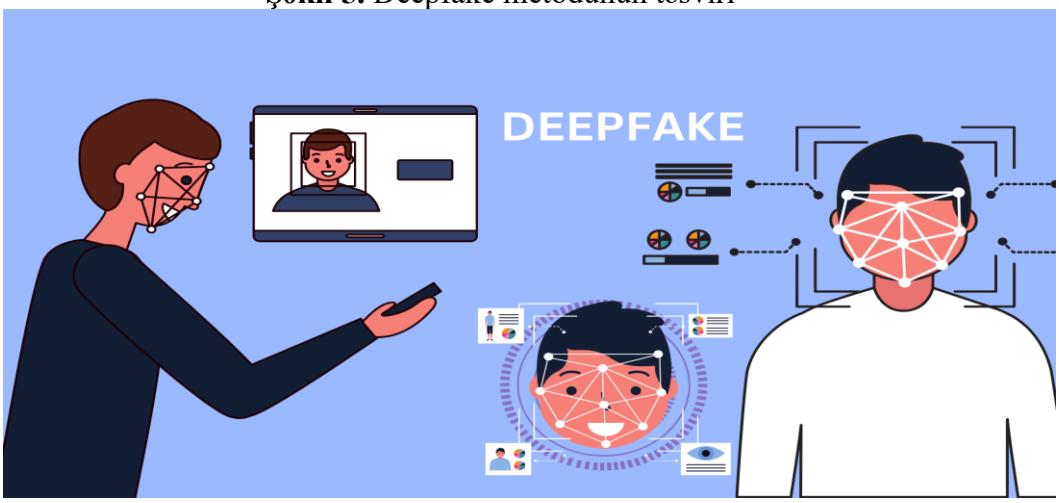
Mənbə. Müəllif tərəfindən tərtib edildi

Sosial mühəndisliyin qarşısını istifadəçilərin təhlükəsizlik şüurunu və şüurlu davranışını artırmaqla almaq mümkündür. Ona görə də istifadəçilər etibarlı mənbələrdən gələn məlumat və sorğulardan xəbərdar olmalı və şəxsi məlumatlarını paylaşmazdan əvvəl diqqətlə düşünməlidirlər. Bundan əlavə, təşkilatlar işçilərini

daha çox məlumatlandırmaq və sosial mühəndislik hücumlarına hazır olmaq üçün müntəzəm olaraq təhlükəsizlik təlimləri keçirə bilər.

Deepfake nədir? - "Deepfake"-lər saxta, lakin dərin öyrənmə alqoritmlərindən istifadə etməklə yaradılmış real video və ya qeydlərdir. "Deepfake" texnologiyası real insanın imicinə və ya səsinə üz ifadələri, səs tonu və hərəkət kimi faktorlar əlavə edərək, özünü başqa insan kimi göstərən saxta videolar yarada bilir.

Şəkil 5. Deepfake metodunun təsviri



Mənbə. Müəllif tərəfindən tərtib edildi

Bu texnologiyadan sui-istifadə siyasi manipulyasiya, şantaj, saxta xəbərlər, onlayn təqib və digər cinayətlər üçün alət kimi istifadə oluna bilər. Deepfake videoları real videolara bənzədiyi üçün cəmiyyətdə inamın itirilməsinə səbəb ola bilər.

Buna görə də, dərin saxta videoları aşkar etmək və aşkar etmək üçün avtomatik tanınma alqoritmləri, blokçeyn əsaslı texnologiyalar və digər üsullardan istifadə etməklə deepfake texnologiyası ilə mübarizə aparmaq üçün səylər göstərilir. Bundan əlavə, deepfake texnologiyasından sui-istifadənin qarşısını almaq üçün qanunlar və qaydalar hazırlanır.

“Samsung” şirkətinin Moskvada yerləşən Süni İntellekt Mərkəzi tərəfindən tək üz şəkli və yaxud, rəsmdən yüksək keyfiyyətli saxta video əldə etmək üçün texnologiya hazırladığı açıqlanıb.

Mona Lizanın mimikaları "Deepfake" texnologiyası sayəsində hərəkətə gətirilərkən tədqiqatçılar Merlin Monroe, Albert Eynşteyn və Fyodor Dostoyevski kimi məşhur adların, eləcə də Mona Liza tablosunun videolarını çəkərək YouTube-da yerləşdiriblər. [1]

Kriptoqrafik hücumlar: Kriptoqrafik hücum, şifrələnmiş məlumatın şifrəsini açmaq və ya şifrələnmiş sistemi zəiflətmək üçün şifrələmə alqoritmi və ya protokolundan istifadə edən hücumdur. Bu cür hücumlar şifrələmə üsullarının sındırıla biləcəyi və ya zəifliklərin aşkar edildiyi vəziyyətlərdə həyata keçirilir.

Şəkil 6. Kriptoqrafik hücumların təsviri

Mənbə. Müəllif tərəfindən tərtib edildi.

Bəzi kriptoqrafik hücum növləri və nümunələri bunlardır:

Kobud güc hücumu : Bu hücumda bütün parol kombinasiyaları sınaqdan keçirildiyi üçün şifrələmə açarını tapmaq çox vaxt aparır.

Lüğət hücumu: Bu hücumda söz siyahısından istifadə edərək şifrələməni sındırmağa cəhd edilir. Word siyahıları istifadəçinin parolunu tapmaq ehtimalını artırır.

Meet-in-the-middle hücumu: Ortada qarşılanma hücumu mənasını verən bu tip hücumda şifrələmə və şifrənin açılması üçün istifadə edilən şifrələmə açarını tapmaq üçün kriptoqrafik alqoritmin ortada qarşılanmasından istifadə edilir.

Yan kanal hücumu: Bu tip hücumda şifrələmə açarı və ya parol şifrələmə prosesində istifadə olunan aparatdakı zəiflikdən istifadə etməklə əldə edilir.

Xarici mesaj hücumu: Bu tip hücumda təcavüzkar şifrələnmiş mesajın içərisinə mesaj yerləşdirir və orijinal mesajı əldə etmək üçün şifrələmə alqoritmindəki zəiflikdən istifadə etməyə çalışır.

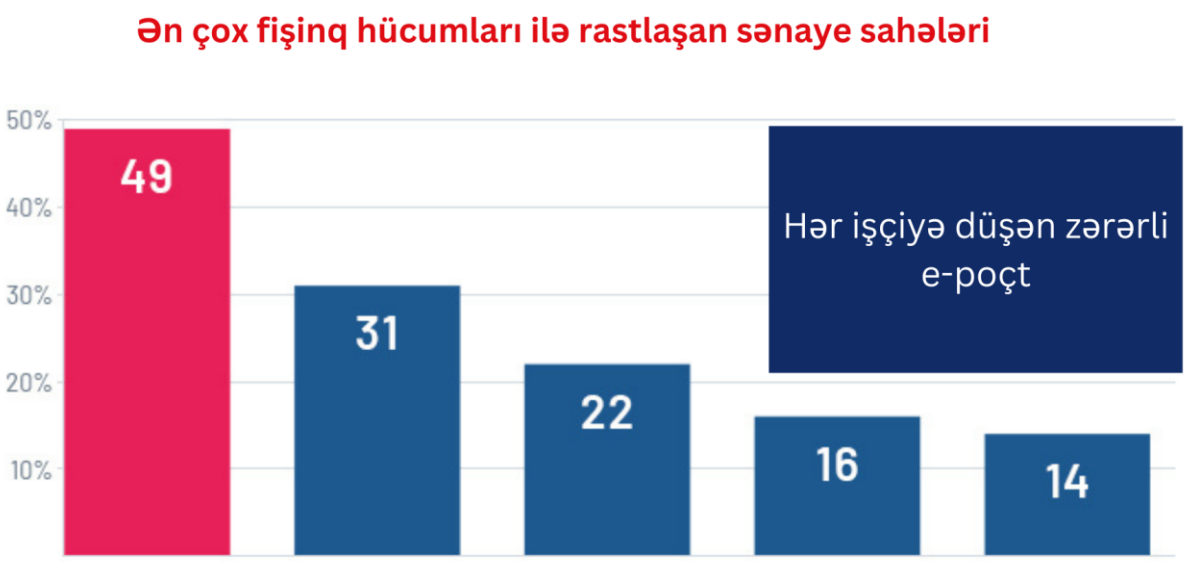
Kriptoanaliz: Bu tip hücumda şifrəli mətnin təhlili ilə şifrələnmiş açarlar və ya mesajlar əldə edilməyə çalışılır.

Şifrələmə hücumlarının qarşısını almaq üçün təhlükəsiz və mürəkkəb şifrələmə alqoritmindən istifadə etmək və şifrələmə açarlarını mütəmadi olaraq dəyişmək tövsiyə olunur olaraq. Bundan əlavə, şifrələmə açarları və ya şifrələnmiş mesajlar təhlükəsiz mühitdə ötürülməli və saxlanmalıdır.

Ümumi statistik araşdırmalara görə nəticə çıxarsaq, fişinq e-poçtlarına ən çox rast gəlinən təşkilatlardan nümunələr aşağıdakı şəkildə öz əksini tapmışdır.

Burada ən çox fişinq hücumları ilə qarşılaşan pərakəndə obyektləridir ki, bu da 49% təşkil etməkdədir. İkinci sırada istehsal sahələri dayanır ki, bu da 31% təşkil edir. Qida və işki müəssisləri 22%-lə bu sırada üçüncü yerdədir. Dördüncü yerdə isə 16%-lə tədqiqat və inkişaf sahələri qərarlaşmaqdadır. Ən sonuncu və digər sahələrə nəzərən daha az fişinqlə qarşılaşan sahələrə isə texnologiya sahələri daxil edilir.

Şəkil 7. Fişinq hücumlarına rast gəlinən sənaye sahələrinin statistik göstəricisi.



Mənbə. Müəllif tərəfindən tərtib edildi

II FƏSİL. MÜASİR SÜNİ İNTELLEKT METOD VƏ TEXNİKALARININ PERFORMANS ANALİZİ

2.1 Ekspert sistemlərin tədqiqi

Ekspert sistemi müəyyən bir sahədə insan təcrübəsinə bənzər qərarlar qəbul etmək və problemləri həll etmək üçün süni intellektin tətbiqidir. Bu sistemlər bilik bazaları, nəticə çıxarma mühərrikləri və istifadəçi interfeysləri kimi komponentlərdən istifadə edərək problemləri təhlil edir və həllər təklif edir.

Ekspert sisteminin dəqiqliyini və etibarlılığını təmin etmək üçün bir sıra sınaq və yoxlama prosesləri tələb olunur. Bu proseslər adətən sistemin dizayn mərhələsində həyata keçirilir və sistem sınaq nəticələrinə uyğun olaraq hazırlanır.

Ekspert sistemləri real dünya problemlərinə tətbiq edildikdə, nəticələrin dəqiqliyi və etibarlılığı təsdiqləmə prosesi ilə təmin edilir. Bu proseslər sistemin real vaxt rejimində istifadəsi zamanı qəbul edilən qərarların düzgünlüyünü qiymətləndirmək üçün istifadə olunur.

Ekspert sistemləri bir çox sahədə istifadə edilə bilər. Məsələn, tibb sahəsində onlar həkimlərə diaqnoz və müalicə tövsiyələri ilə bağlı qərar qəbul etməyə kömək edə bilərlər. Onlar həmçinin maliyyə, mühəndislik, istehsalat və digər sənaye sahələrində istifadə edilə bilər. Ekspert sistemindən istifadə insan səhvlərini azaltmağa və məhsuldarlığı artırmağa kömək edir. Bununla belə, dəqiqliyi təmin etmək üçün müvafiq sınaq və yoxlama prosedurları həyata keçirilməlidir.

Nəticə olaraq, ekspert sistemi müəyyən bir sahədə insan təcrübəsinə bənzər qərarlar qəbul etmək üçün süni intellektin tətbiqidir. Bu sistemlərin düzgünlüyünü və etibarlılığını təmin etmək üçün sınaq və yoxlama prosesləri tələb olunur. Ekspert sistemindən istifadə insan səhvlərini azaltmağa və məhsuldarlığı artırmağa kömək edir.

Bununla belə, mütəxəssislərin kompüterlərdən istifadə edərək həll edə biləcəyi mürəkkəb problemləri həll etməyə imkan verən sistemlər ekspert sistemləri adlanır. Ekspert sistemləri süni intellekt proqramlaşdırmasıdır. Onlar adətən alqoritmlərdən istifadə etmirlər və ən əsası verilənlər bazalarından daha çox məlumata malikdirlər. Ekspert sistemində simulyasiya və hesablama kimi funksiyalar ola bilər. Bu sistemlərin çoxlu müxtəlif modelləri var və ən çox yayılmışı qayda əsaslı modeldir. Qayda əsaslı sistem - sistem trafikini yoxlayır və hücumun aşkarlanması zamanı müəyyən edilmiş qaydalara uyğun olaraq davranışı təsnif etmək üçün qaydalar yaradır. Bu sistemlər təhlükəsizlik səviyyələrinin seçilməsini xeyli asanlaşdırır və məhdud resurslardan yüksək istifadə üçün təlimat verir. Ekspert sistemlərindən istifadə etməklə müdaxilənin aşkarlanması mümkündür.

2.2 Ağıllı agentlərin (Intelligent agents) tətbiqi

Ağıllı agentlər insanların etdiklərini etmək üçün proqramlaşdırılmış süni intellekt sistemləridir. Bu sistemlər xüsusi tapşırıqları yerinə yetirmək üçün əvvəlcədən proqramlaşdırılmış və ya öyrənilmiş alqoritmlərdən istifadə edərək məlumatları emal edir. Son illərdə bu texnologiyanın potensialı getdikcə daha çox tanınır və bir çox sənaye sahələrində tətbiq olunur. Ağıllı agentlər insanların qərar vermə proseslərində süni zəka və öyrənmə alqoritmlərini yerinə yetirə və istifadə edə biləcəkləri bir çox işi yerinə yetirə bilən proqram və sistemlərdir. Bu agentlər tətbiqin bir çox müxtəlif sahələrində istifadə oluna bilər.

Məsələn, ağıllı agentlər təhlükəsizlik tətbiqlərində istifadə oluna bilər. Təhlükəsizlik agenti binanın təhlükəsizliyini təmin etmək üçün girişlərə nəzarət edə və şübhəli hərəkətləri aşkar edə bilər. Eynilə, bankların fırıldaqçılığın qarşısını almaq üçün istifadə etdiyi fırıldaqçılıq aşkarlama sistemlərində də ağıllı agentlərdən istifadə etmək olar.

Ağıllı agentlər istifadəçinin əmrləri ilə interaktiv fəaliyyət göstərən “davranış” agentləridir. İstifadəçi ilə təkbətək əlaqə qurur və ondan gələn əmrlərə uyğun olaraq davranışlarını tənzimləyirlər. Ağıllı agent əvvəlcədən mövcud məlumat əsasında qərarlar qəbul etmək və performans meyarlarını inkişaf etdirmək mexanizminə malikdir. O, müdafiə sistemləri ilə əməkdaşlığı simulyasiya etmək üçün ağıllı agentlərdən istifadə edərək Paylanmış Xidmətdən imtina (DDoS) hücumlarından müdafiə edə bilər.

Ağıllı agentlərin tətbiqi bir çox sahələrdə əhəmiyyətli təsir göstərə bilər. Məsələn, bir şirkətin müştəri xidməti şöbəsi müştərilərin suallarını cavablandırmaq üçün ağıllı agentlərdən istifadə edərək, cavabları tez və dəqiq verə bilər. Ağıllı agentlər səhiyyə sənayesində də istifadə edilə bilər. Xəstəxanada tibbi diaqnostika və müalicə prosesində ağıllı agentlərdən istifadə etməklə həkimlər tez bir zamanda diaqnoz qoya və müalicə tövsiyələri verə bilərlər.Son olaraq, ağıllı agentlər tibb sahəsində istifadə oluna bilər. Tibbi görüntü sistemlərində istifadə olunan ağıllı agentlər tibbi görüntüləri analiz edərək xəstəliklərin aşkarlanmasına kömək edə bilərlər. Bundan əlavə, ağıllı agentlərdən istifadə ilə tibbi tədqiqatlarda istifadə olunan məlumatları təhlil etmək və müalicə üsullarının təkmilləşdirilməsinə kömək etmək olar.

Bundan başqa, onlardan ağıllı agentlər, ağıllı ev sistemləri kimi ağıllı cihazlarda da istifadə etmək olar. Bu cihazlar evin avtomatlaşdırılması və ev sahibinin təhlükəsizliyi üçün işləyə bilər.

Ağıllı agentlərdən sənaye istehsalı proseslərində də istifadə etmək olar. İstehsal xəttindəki robotlar məhsulları avtomatik sınaqdan keçirə və keyfiyyətə nəzarət proseslərini həyata keçirə bilərlər.

Ağıllı agentlərin tətbiqi müdafiə sənayesində də böyük rol oynaya bilər. Məsələn, hərbçilər hücumları aşkar edərək və müdafiə tədbirləri görərək tez və effektiv cavab verə bilirlər.

Texnologiyadan istifadə də etik suallar doğura bilər. Məsələn, şəxsi məlumatların məxfiliyi kimi məsələlər ağıllı vasitəçilərdən istifadə edərkən diqqət yetirilməli olan mühüm məsələlərdəndir. Bu problemlərin həlli ağıllı agentlərin effektiv istifadəsi üçün çox vacibdir.

2.3 Genetik alqoritmlərin işlənməsi

Genetik alqoritm bioloji təkamül prosesində təbii seçmə prinsipinə əsaslanan optimallaşdırma üsuludur. Bu alqoritmlər populyasiyalar arasında genetik krossover və mutasiya əməliyyatlarını yerinə yetirərək optimal nəticələr əldə etmək üçün nəzərdə tutulmuşdur.

Genetik alqoritmlərdən istifadə sənayenin bir çox sahələrində geniş istifadə olunur. Məsələn, genetik alqoritmlər məhsul dizaynında və istehsalda optimallaşdırmada istifadə edilə bilər. Bundan əlavə, maliyyə sahəsində də istifadə olunan bu alqoritmlər investisiya strategiyalarında istifadə oluna bilər və portfelin optimallaşdırılmasını təmin edə bilər.

Genetik alqoritmin emalı çox mürəkkəb bir prosesdir. Bu alqoritmlər spesifik populyasiyaları təmsil edən genlərdən ibarətdir.

Ən yaxşı nəticələr verən genləri müəyyən etmək üçün proses bir neçə təkrarlamadan keçdi. Ancaq bu, çox vaxt aparan və hesablama baxımından intensiv bir prosesdir. Yanlış aparılarsa, genetik alqoritmlərin inkişafı səhv nəticələr verə və yanlış qərarlara səbəb ola bilər.

Yekun bir nəticəyə gəlmiş olsaq, genetik alqoritm emalı bir çox sənaye sahələrində istifadə olunan çox faydalı optimallaşdırma üsuludur. Lakin bu texnologiyadan düzgün istifadə etmək üçün kifayət qədər hesablama gücünə malik olmaq və prosesi düzgün idarə etmək vacibdir.

III FƏSİL. ZORLA MÜDAXİLƏNİN AŞKARLANMASI SİSTEMLƏRİ (IDS) VƏ ONLARIN QARŞISININ ALINMASININ SÜNİ İNTELLEKT METODU İLƏ İŞLƏNMƏSİ

3.1 Süni intellektə yönəlik hücum növlərinin tədqiqi

Süni intellekt (AI) və ya ingiliscə terminoloji adıyla desək, "Artifical Intelligence"(AI) sistemləri inkişaf etdikcə və onun köməyi ilə bir çox əməliyyatlar çevik həll edilməyə başladıqca süni intellekt sistemlərinin özləri də hücumlara məruz qala bilər.

Ən müasir ***süni intellekt*** **(AI)** sistemlərinin arxasında duran üsullar belə sistematik olaraq "AI hücumları" kimi tanınan kibertəhlükəsizlik hücumlarının yeni sinfinə qarşı həssasdır. Bu cür hücumlar təcavüzkarlara öz zərərli son məqsədlərinə çatmaq və davranışlarını dəyişdirmək üçün bu sistemləri manipulyasiya etməyə imkan verir. Süni intellekt sistemləri cəmiyyətin kritik hissələrinə getdikcə daha çox inteqrasiya etdir ki, bu da süni intellekt hücumları riski ilə milli təhlükəsizlik üçün əhəmiyyətli təsirləri ola biləcək yeni sistem zəifliklərini təmsil edir. Bu "AI hücumu" ənənəvi kiberhücumlardan əsaslı şəkildə fərqlənir.

Kod "baqları" və ya insan səhvi ilə idarə olunan ənənəvi kiberhücumlardan fərqli olaraq, AI hücumları hazırda həll edilə bilməyən əsas AI alqoritmlərinin xas məhdudiyyətləri ilə mümkün olur. Bundan əlavə, AI hücumları kiberhücumları həyata keçirmək üçün istifadə edilə bilən obyektlər dəstini əsaslı şəkildə genişləndirir.

İlk dəfə olaraq, fiziki obyektlər artıq kiberhücumlar üçün istifadə edilə bilər (məsələn, dayanma nişanına bir az lent yapışdırmaq kifayətdir və süni intellekt hücumu dayanma nişanını özü idarə edən avtomobilin gözündə yaşıl işığa çevirə

bilər). Məlumatların silahlanması dedikdə məlumatların toplanması, saxlanması və istifadə edilməsi qaydasında dəyişikliklər tələb edir.

Cəmiyyətin həssas olduğu vacib nöqtələrə nəzər yetirsək, beş sektor süni intellekt hücumlarından ən çox təsirlənə bilər: məzmunun filtrasiyası, hərbi sistem, hüquq-mühafizə orqanları, süni intellektlə əvəzlənən ənənəvi insan vəzifələri və vətəndaş cəmiyyəti. Bu domenlər cəlbedici hücum hədəfləridir və AI getdikcə kritik məqsədlər üçün istifadə edildiyi üçün getdikcə daha həssas olur.

Hesabat süni intellekt hücumlarının qarşısını almaq üçün "AI təhlükəsizliyinə uyğunluq" proqramını təklif edir. "AI Təhlükəsizliyə Uyğunluq" proqramının dövlət siyasətinin inkişafı süni intellekt sistemlərinin hücuma məruz qalma riskini və uğurlu hücumun təsirini azaldacaqdır.

Uyğunluq proqramları buna maraqlı tərəfləri süni intellekt sistemlərinin yerləşdirilməsi zamanı hücum risklərinin və səth sahələrinin nəzərə alınması, hücumların həyata keçirilməsini çətinləşdirmək üçün İT islahatlarının qəbul edilməsi və hücuma cavab tədbirlərinin hazırlanması daxil olmaqla sistemləri AI hücumlarından qorumaq üçün bir sıra ən yaxşı təcrübələri qəbul etməyə təşviq etməklə nail olacaq. Tənzimləyicilər hökumətlərdən qaydalara riayət etməyi və yüksək riskli AI-dən istifadə etməyi tələb etməlidir.

Tənzimləyicilər süni intellekt sistemlərinin hökumətlərə istifadəsi və satışı ilə bağlı hökumət qaydalarına riayət etməyi tələb etməlidir. Süni intellekt hücumlarının ciddi ictimai nəticələrinin qarşısını almaq üçün özəl sektordakı tənzimləyicilər yüksək riskli tətbiqlər üçün uyğunluğu təmin etməli, eyni zamanda aşağı riskli tətbiqləri istəyə bağlı etməlidir. Bu, yeniliyin maneəsiz davam etməsinə imkan verir.

Müasir dövrdə terrorçulara məqsədlərinə çatmaq üçün mürəkkəb silahlara ehtiyac yoxdur. Bunun üçün sadəcə bir rulon elektrik lenti və etibarlı bir cüt idman ayaqqabısı lazımdır. Yol kəsişmələrində dayanma nişanları üçün xırda lent parçalarını gizlicə yapışdırmaqla, onlar heç bir çətinlik çəkmədən nişanların siqnallarını dəyişdirə, onları özü idarə edən nəqliyyat vasitələri üçün sayrışan yaşıl işıqlara çevirə bilərlər. Bu cür zarafatlar az tıxac olan kəsişmələrdə qəzalara səbəb ola bilər. Bununla belə, səs-küylü bir metropolda sıx və vacib kəsişmələrdə yerləşdirilsə, onların hərəkətləri şəhərin nəqliyyat infrastrukturunu şikəst edə bilər. Aydındır ki, bir lentə iki dollardan az investisiya qoymaq belə böyük gəlirlər əldə edə bilər.

Kompüter elmində kiçik bir qrup insan süni intellektlə bağlı problemləri həll etmək üçün çox çalışır Bu xüsusi çətinlik qaranlıq olsa da, yaxın gələcəkdə qlobal iqtisadi, hərbi və sosial təminata güzəştə getmək potensialına malikdir.

Təəssüf ki, süni intellektdə istifadə olunan alqoritmlərin öyrənmə prosesi zamanı onları hücumçulara qarşı həssas edən zəifliyi var. Yenə də həll yoluna ümid var. Dayanma işarəsi vandalizm kimi görünəndə AI sistemi yaşıl işıq görür. “AI hücumları” kimi tanınan süni intellektə qarşı bu cür hücumlar getdikcə genişlənir. Nə qədər təhlükəli olsalar da, qabaqcıl AI metodlarının xas məhdudiyyətlərindən istifadə edən bir çox hücumlar var. Bu zəiflik bu üsulların üzləşdiyi narahatedici problemdir. Müəyyən taktikalardan istifadə etməklə fərdlər qabaqcıl AI texnologiyalarını asanlıqla poza bilər.

Bu üsullar kiçik bir lent parçasının dayanma işarəsi üzərində yapışdırılmasından tutmuş, insan tərəfindən diqqətdən kənarda qalacaq görüntüyə görünməz rəqəmsal zibil səpələnməsinə qədər dəyişir. Bundan əlavə, AI sistemlərini çirkləndirmək və müəyyən vaxtlarda və yerlərdə təxribata imkan verən gizli portallar quraşdırmaq potensialı var.

Təhlükələr artıq mövcuddur, nümunələrə kəşfiyyat missiyaları zamanı dronları aldatmaq, terror təbliğatı üçün məzmun filtrlərini pozmaq, avtomobil qəzasına səbəb olan qırmızı işıqda yanmaq və s. daxildir. Bu, hamımıza təsir edən geniş və ciddi problemdir.

Süni intellekt avtoritar hökumətlər tərəfindən nəzarət üçün istifadə edildiyi üçün bəzi AI hücumları pis bir şey kimi görünməyə bilər. Süni intellekt hücumları VPN və Torun oynadığı rola bənzər hökumət zülmünə qarşı dayaq rolunu oynaya bilər. Bununla belə, qeyd etmək lazımdır ki, bütün AI tətbiqləri yaxşı deyil.

İstifadəsindən asılı olmayaraq, süni intellekt hücumları son vaxtlar başlıqlarda üstünlük təşkil edən kibertəhlükəsizlik problemlərindən fərqlidir. Bu hücumlar düzəldilə bilən kod xətaları deyil - onlar AI alqoritmlərinin nüvəsinə xasdır. Buna görə də, bu süni intellekt zəifliklərindən istifadə etmək hədəf sistemi "sındırmaq" tələb etmir. Əslində, bu kritik sistemlərə hücum etmək həmişə kompüter tələb etmir.

Bu, hökumətlərin və bizneslərin bir araya gətirdiyi mövcud kibertəhlükəsizlik və siyasət alətləri ilə həll edilə bilməyən yeni kibertəhlükəsizlik problemləri toplusudur və eyni zamanda, problemin həlli yeni yanaşmalar və həll yolları tələb edir. Zaman keçdikcə tədqiqatçılar bu problemlərin bəzilərinin texnoloji yollarını kəşf edə bilərlər.

Burada əsas məqsəd siyasətçiləri, sənaye liderlərini və kibertəhlükəsizlik ictimaiyyətini ortaya çıxan bu problem haqqında məlumatlandırmaq, cəmiyyətin hansı sahələrinin daha çox risk altında olduğunu müəyyənləşdirmək və tapılan bu mühüm yeni dövrdə təhlükəsiz qalmaq üçün nə edilə biləcəyini təsvir etməkdir.

Hesabat dörd bölməyə bölünür. Birincisi, mövcud AI sistemlərinə necə hücum edildiyinin, bu hücumların mövcud olduğu formaların və onların taksonomiyasının onları necə təsnif etdiyinin başa düşülən və hərtərəfli təsviridir.

İkincisi, hesabat bu yeni növ zəifliyin təsir etdiyi ən kritik sahələri müəyyən edir. Bu yeni təhlükənin təsirinə məruz qalan sistemlərin sayı yalnız süni intellektin müasir dünyada geniş vüsət alması ilə artsa da, bu hesabat dərhal diqqət tələb edən beş yüksək prioritet sahəyə diqqət yetirir:

-məzmunun filtrasiyası,

-hərbi,

-qanunla yerinə yetirilən insan vəzifələrinin dəyişdirilməsi,

-icra,

-süni intellekt,

-vətəndaş cəmiyyəti.

Üçüncüsü, süni intellekt zəifliklərini daha geniş kibertəhlükəsizlik kontekstində yerləşdirir. Süni intellekt hücumlarının təbiətcə fərqli olan və mövcud kibertəhlükəsizlik boşluqlarına cavab tələb edən yeni hücum şaqulisini təmsil etdiyi iddia edilir. Bu bölmə həmçinin AI hücumlarının hücum kiber silahları kimi istifadəsini təsvir edir.

Süni intellekt hücumlarının qarşısını almaq üçün "AI təhlükəsizlik uyğunluğu" proqramı ideyasını təklif edir. Bu uyğunluq proqramları AI sistemlərinə hücum riskini azaldır və uğurlu hücumların təsirini azaldır. Onlar bunu maraqlı tərəfləri sistemləri AI hücumlarından qorumaq üçün ən yaxşı təcrübələr toplusunu qəbul etməyə, o cümlədən AI sistemlərini yerləşdirərkən hücum risklərini və səthləri nəzərə almaq, hücumların həyata keçirilməsini çətinləşdirən İT islahatlarını qəbul etmək və hücuma cavab planı yaratmaqla edəcəklər.

Bəs hücumun zərərini necə azaltmaq lazımdır?

Tənzimləyicilər dövlət və özəl sektorların bəzi hissələrində uyğunluq tələb etməlidirlər. Hökumətin süni intellektdən istifadəsinə uyğunluq dövlət sektorunda məcburi olmalıdır və hökumətlərə AI sistemləri satan özəl şirkətlər üçün tələb olmalıdır. Özəl sektorda, bu sürətlə dəyişən sahədə innovasiyaları pozmamaq üçün yüksək riskli özəl sektorun AI tətbiqləri üçün uyğunluq məcburi, lakin aşağı riskli olanlar üçün istəyə bağlı olmalıdır. Bu siyasət süni intellekt hücumları qarşısında icmanın, ordunun və iqtisadiyyatın təhlükəsizliyini yaxşılaşdıracaq. Lakin həm siyasətçilər, həm də maraqlı tərəflər üçün bu təhlükəsizliyin həyata keçirilməsi istiqamətində ilk addım bizim diqqətimizi hazırda yönəltdiyimiz problemi anlamaqdan başlayır.

Süni intellekt sistemlərinin niyə eyni zəifliklərə meylli olduğunu başa düşmək üçün biz AI alqoritmlərinin, daha dəqiq desək, onların istifadə etdiyi maşın öyrənmə üsullarının necə “öyrəndiyini” qısaca araşdırmalıyıq. Kəşfiyyatçı kimi, süni intellekt sistemlərini gücləndirən maşın öyrənmə alqoritmləri verilənlərdən nümunələr çıxararaq "öyrənirlər".

Bu nümunələr, təsvirdə mövcud olan obyektlər kimi, tapşırığa uyğun olan yüksək səviyyəli anlayışlarla əlaqələndirilir. Məsələn, özünü idarə edən avtomobildə dayanma əlamətlərini tanımaq üçün AI alqoritmini öyrənmək tapşırığını nəzərdən keçirək. Bu tapşırığın yerinə yetirilməsi üçün alqoritm ona dayanma işarələrinin yüzlərlə və ya minlərlə nümunələrindən ibarət verilənlər toplusunu göstərməklə və onları təmsil edən rəng və forma nümunələrini çıxararaq “öyrənir”. Əgər alqoritm sonradan müəyyən işarənin dayanma işarəsi olub-olmadığını müəyyən etmək istəsə, təsviri skan edir və dayanma işarələri ilə əlaqələndirməyi öyrəndiyi nümunələri axtarır. Nümunə uyğun gələrsə, alqoritm avtomobilə dayanmağı söyləyə bilər.

Alqoritm ekvivalent olaraq, model başqa bir simvolun, məsələn, yeni daha sürətli sürət həddi ilə uyğunlaşarsa, avtomobilə sürətləndirməyi əmr edə bilər.

FUSAG almanları aldatmaq üçün şişirdilmiş şarlar üzərində hansı naxışların çəkilməli olduğunu məharətlə ixtira edə bildiyi kimi, düşmənlər də süni intellekt sistemini aldadacaq hədəflər üzərində dəyişdirilmiş nümunələr yaratmaq üçün "giriş hücumu" kimi tanınan AI hücumundan istifadə edə bilərdilər. səhvlər etmək. Bu hücum mümkündür, çünki hədəfdəki nümunələr verilənlər bazasındakı dəyişikliklərə uyğun gəlmədikdə sistem ixtiyari nəticələr verir, təcavüzkarlar qəsdən bu uyğunsuz nümunələri əlavə etdikdə olduğu kimi. Bununla belə, tank nümunəsindən fərqli olaraq, bu naxışlar və ya işarələr o qədər də görkəmli olmamalıdır. Çünki süni intellekt alqoritmləri informasiyanı insanlardan fərqli şəkildə emal edir.[2]

Beləliklə, insanı və ya süni intellekt sistemini aldatmaq üçün şarın tanka bənzədilməsi lazım ola bilsə də, süni intellekt sistemini məhv etmək üçün sadəcə bir neçə alov və ya təsvirdə bir neçə pikselə kiçik dəyişiklik lazımdır. Bu giriş hücumları AI hücumunun yalnız bir növüdür.

Zəhərlənmə hücumu kimi tanınan digəri süni intellekt sisteminin müəyyən şərtlərdə düzgün işləməsinə mane ola bilər və ya hətta sonradan düşmən tərəfindən istifadə edilə bilən arxa qapı daxil edə bilər. Bənzətmə ilə davam etsək, zəhərli hücum hipnozla eyni olardı və alman analitikləri hər dəfə müttəfiqlərə zərər vermək üçün istifadə oluna biləcək qiymətli məlumatları görmək ərəfəsində olanda gözlərini yummağa məcbur edərdi.

Ümumiyyətlə, bu hücumlar ciddi kibertəhlükələrin xüsusiyyətlərinə malikdir Fiziki hədəflərdə nöqtələr və ya əyri xətlər şəklində görünə bilər və ya süni intellekt sistemlərinin DNT-sində gizlənə bilərlər.

- Hücum insan gözü üçün tamamilə görünməz ola bilər. Bunun əvəzinə, onlar gözdən gizlənə bilər, bu da onları ətrafları ilə mükəmməl birləşmiş kimi göstərir.

Bəs AI hücumu tam olaraq nədir? Niyə mövcuddur? onlar necə görünürlər Bu suallara cavab vermək üçün indi bu hücumların texniki əsaslarını anlamağa müraciət edirik.

Süni intellekt (AI) hücumu, süni intellekt sisteminin nasazlığına səbəb olmaq məqsədi ilə qəsdən manipulyasiya edilməsidir. Bu hücumlar əsas alqoritmlərdə müxtəlif zəiflikləri hədəf alaraq müxtəlif formalarda ola bilər:

- Giriş və yaxud daxiletmə hücumu: Təcavüzkarın məqsədlərinə xidmət etmək üçün sistemin çıxışını dəyişdirmək üçün AI sisteminə girişin məzmununu manipulyasiya etməkdir. Çünki hər bir süni intellekt sisteminin mərkəzində sadə bir maşın dayanır – o, daxilolman; qəbul edir, müəyyən hesablamalar həyata keçirir və çıxışı qaytarır – girişlə manipulyasiya etmək təcavüzkarın sistemin çıxışını manipulyasiya etməyə imkan verir.

- Zəhərləmə hücumu: Süni intellekt sisteminin yaradılması prosesinin alt-üst edilməsi, nəticədə ortaya çıxan sistemin təcavüzkarın nəzərdə tutduğu şəkildə işləməməsinə səbəb olur. Zəhərlənmə hücumunu həyata keçirməyin sadə bir yolu prosesdə istifadə olunan məlumatları məhv etməkdir. Bunun səbəbi süni intellektə güc verən ən qabaqcıl maşın öyrənmə üsullarının tapşırığı necə yerinə yetirməyi "öyrənməklə" işləməsidir. Lakin onlar yalnız bir mənbədən və yalnız bir məlumatdan "öyrənirlər". Məlumat su, yemək, hava kimidir. Zəhərlənmə hücumları da öyrənmə prosesinin özünə müdaxilə edə bilər.

Süni intellekt sistemləri kritik kommersiya və hərbi tətbiqlərdə yerləşdikcə, bu hücumlar ciddi və hətta ölümcül nəticələrə səbəb ola bilər. AI hücumları zərərli son məqsədlərə nail olmaq üçün bir neçə yolla yerləşdirilə bilər:

- Zərər vurmaq: Təcavüzkar AI sistemini sıradan çıxararaq zərər vermək istəyir. Buna misal olaraq avtonom avtomobili aldadaraq dayanma işarəsinə məhəl qoymayan hücumu göstərmək olar.

Təcavüzkar süni intellekt sisteminə hücum edərək, bir dayanacaq nişanını başqa bir işarə və ya simvol kimi səhv müəyyənləşdirə bilər, özü idarə edən nəqliyyat vasitələrinin dayanma nişanlarına məhəl qoymamasına və digər nəqliyyat vasitələri və piyadalarla toqquşmasına səbəb ola bilər.

- Nədənsə gizlənmə: Təcavüzkar AI sistemi tərəfindən aşkarlanmamaq istəyir. Buna misal olaraq sosial şəbəkələrdə terror təbliğatının yerləşdirilməsinin qarşısını almaq üçün nəzərdə tutulmuş məzmun filtrinin nasaz işləməsi və materialın sərbəst şəkildə yayılmasına imkan verən hücumu göstərmək olar.

- Sistemə etimadın qırılması: Təcavüzkarlar operatorların süni intellekt sisteminə inamını itirməsini istəyirlər ki, bu da sistemin bağlanmasına səbəb olur. Buna misal olaraq, avtomatlaşdırılmış təhlükəsizlik xəbərdarlıqlarının təkrarlanan hadisələri səhv olaraq təhlükəsizlik təhdidləri kimi təsnif etməsinə səbəb olan və sistemləri oflayn vəziyyətə salan çoxlu sayda yanlış pozitivlərə səbəb olan hücumdur. Yoldan keçən sahibsiz pişiyi və ya yıxılmış ağacı təhlükəsizlik kimi təsnif etmək üçün video əsaslı təhlükəsizlik sisteminə hücum etmək təhlükəsizlik sistemini oflayn vəziyyətə sala bilər ki, bu da öz növbəsində real təhlükələrin aşkarlanmasından yayınmağa imkan verir.

Süni intellektin son on ildə görünməmiş uğurunu nəzərə alsaq, bu hücumların mümkün olduğunu və onların hələ də düzəldilmədiyini öyrənmək daha təəccüblüdür. İndi bu hücumların niyə mövcud olduğuna və onların qarşısını almaq niyə bu qədər çətin olduğundan danışaq.

Niyə süni intellekt hücumları var? Sİ hücumları, təcavüzkarın sistemin uğursuzluğuna səbəb olmaq üçün istifadə edə biləcəyi Sİ alqoritmlərinin əsas məhdudiyyətləri səbəbindən mövcuddur. Ənənəvi kibertəhlükəsizlik hücumlarından fərqli olaraq, bu zəifliklər proqramçı və ya istifadəçi səhvinin nəticəsi deyil. Bunlar, sadəcə olaraq, müasir metodların çatışmazlıqlarıdır. Daha açıq desək, süni intellekt sistemlərinin yaxşı işləməsini təmin edən alqoritmlər mükəmməl deyil və onların sistemli məhdudiyyətləri rəqiblərin hücumu üçün imkanlar yaradır. Bu, ən azı yaxın gələcək üçün, riyaziyyatda sadəcə həyat faktıdır.

İnsanlar real dünyada bir çox obyekt və ya konsepsiya nümunəsini görərək və öyrəndiklərini daha sonra istifadə etmək üçün beyinlərində saxlayaraq dərk edirlər. Maşın öyrənmə alqoritmləri verilənlər toplusunda bir çox obyekt və ya konsepsiya nümunələrinə baxaraq "öyrənir" və öyrəndiklərini sonradan istifadə üçün modeldə saxlayır. Çox olmasa da, maşın öyrənməsinə əsaslanan süni intellekt tətbiqetmələrində prosesdə heç bir xarici bilik və ya digər sehr istifadə edilmir. O, tamamilə verilənlər bazasına əsaslanır, başqa heç nə.

Süni intellekt hücumlarını başa düşməyin açarı maşın öyrənməsində “öyrənmə” sözünün əslində nə olduğunu və daha da əhəmiyyətlisi nəyin olmadığını anlamaqdır. Unutmayın ki, maşın öyrənməsi verilənlər bazasındakı bir çox konsepsiya və ya obyekt nümunələrinə baxaraq "öyrənir". Daha dəqiq desək, bu nümunələrdə ümumi nümunələri çıxaran və ümumiləşdirən alqoritmlərdən istifadə edir. Bu sxemlər modeldə saxlanılır.

Dayanma işarəsinin aşkarlanması nümunəsindən istifadə edərək öyrənmə alqoritmi nümunə şəklini təşkil edən piksellərdə nümunələri müəyyən edir. Daha sonra modeldən yeni təsvirdə dayanma işarəsini tanıması tələb edildikdə, o, həmin təsvirdə eyni bitmapı axtarır. Dayanma işarələri ilə əlaqələndirməyi öyrəndiyi

nümunəyə uyğun gələn nümunə tapdıqda, dayanma işarəsi tapdığını bildirir. Bunun əvəzinə, yaşıl işıq kimi başqa bir obyektlə uyğunlaşmağı öyrəndiyi naxışa uyğun bir nümunə tapsa, yaşıl işıq tapdığını bildirir. Bu nümunələr yalnız öyrəndikləri nümunələrə əsaslanmamaqla, yeni kontekstlərdə işləməli olduqları üçün "ümumi" olur. Məsələn, yuxarıda qeyd edilən nümunə yalnız məlumat dəstindəkilərə deyil, bütün dayanma işarələrinə uyğun olmalıdır.

Kifayət qədər məlumat nəzərə alınarsa, bu şəkildə öyrənilən nümunələr o qədər yüksək keyfiyyətə malikdir ki, onlar hətta bir çox vəzifələrdə insanları üstələyə bilər. Çünki alqoritm hədəfin bütün müxtəlif təbii təzahürlərinin kifayət qədər nümunəsini görəndə, alqoritm özünü yaxşı aparması üçün lazım olan bütün nümunələri tanımağı öyrənir.

Bununla belə, bu proses artıq nəzərə çarpan bir zəifliyi təqdim edir: o, tamamilə verilənlər bazasından asılıdır. Verilənlər toplusu modelin yeganə bilik mənbəyi olduğundan, bu məlumatlardan öyrənilən model, təcavüzkar tərəfindən zədələnərsə və ya "zəhərlənərsə" pozula bilər.

Təcavüzkarlar modellərin müəyyən nümunələri öyrənməsinin qarşısını almaq üçün verilənlər bazasını zəhərləyə və ya gələcəkdə modelləri aldatmaq üçün istifadə oluna biləcək gizli arxa qapılar quraşdıra bilərlər.

Amma problem bununla bitmir. Hətta pozulmamış verilənlər bazası və yüksək dəqiqlikli modeli fərz etsək belə, bu uğur çox vacib bir xəbərdarlıqla gəlir: mövcud ən müasir maşın öyrənmə modelləri tərəfindən "öyrənilən" nümunələr nisbətən kövrəkdir. Buna görə də, model yalnız öyrənmə prosesində istifadə olunan məlumatlara oxşar məlumatlar üçün uyğundur. Model orijinal verilənlər bazasında gördüyü variasiya tipindən bir qədər fərqli olan verilənlərə tətbiq edildikdə tamamilə uğursuz ola bilər. Bu, təcavüzkarın istifadə edə biləcəyi mühüm məhdudiyyətdir:

Təcavüzkar süni variasiya təqdim etməklə – məsələn, lent parçası və ya digər anormal nümunə – modeli korlaya və təqdim edilən süni model əsasında onun davranışına rəhbərlik edə bilər.

Model yaratmaq üçün istifadə olunan məlumatların miqdarı məhdud olduğundan, təcavüzkarın yarada biləcəyi süni variasiyanın miqdarı qeyri-məhdud olur ki, beləliklə, təcavüzkarın özünəməxsus üstünlüyü olmuş olur.

Bu, dayanma işarəsi lentinə hücumun özünü idarə edən avtomobillərin necə qəzaya uğramasına səbəb ola biləcəyini izah edir. Dayanma işarəsi detektorunu öyrətmək üçün istifadə edilən verilənlər bazası müxtəlif təbii şəraitlərdə dayanma işarəsinin bir çox variantını ehtiva etsə də, bu, təcavüzkarın lent və qraffiti kimi süni şəkildə manipulyasiya edə biləcəyi sonsuz imkanların nümunəsini təqdim etmir. Bu səbəbdən, çox kiçik insan manipulyasiyaları, yaxşı seçildikdə, model tərəfindən öyrənilən nisbətən kövrək nümunələri poza bilər və modelin çıxışına çox böyük təsir göstərə bilər.

Buna görə kiçik bir lent parçası dayanma işarəsini asanlıqla yaşıl işığa çevirə bilər: o, bütün dayanma işarəsini yaşıl işıq kimi göstərməli deyil, sadəcə modelin xüsusi kiçik kövrək modelini aldatmalıdır. Təəssüf ki, bunu etmək asandır.

Bir çox insanlar maşın öyrənmənin belə parlaq çatışmazlıqlara sahib olduğunu öyrəndikdə təəccüblənərdilər. Bunun səbəbi, populyar mədəniyyətin maşın öyrənməsinin əslində insan mənasında "öyrəndiyi" ilə bağlı geniş yayılmış, lakin yanlış bir inam yaratmasıdır. İnsanlar həqiqətən anlayışları və assosiasiyaları öyrənməyi yaxşı bacarırlar. Əgər dayanma nişanı qraffiti və ya kirlə təhrif edilibsə və ya ləkələnibsə, hətta heç vaxt qraffiti və ya çirkli dayanma nişanı görməmiş insanlar belə, hələ də etibarlı və ardıcıl olaraq onu dayanma işarəsi kimi müəyyən edəcəklər və əlbəttə ki, onu tamamilə fərqli obyektlə qarışdırmayacaqlar.

Ancaq indi bilirik ki, indiki AI sistemləri eyni şəkildə işləmir. Hətta dayanma nişanlarını demək olar ki, mükəmməl şəkildə tanıya bilən bir model hələ də dayanma nişanları, hətta insanlar kimi yol nişanları anlayışını bilmir. Sadəcə bilir ki, müəyyən öyrənilmiş nümunələr "dayan işarələri" adlanan etiketlərə uyğun gəlir.

İnsan öyrənməsi ilə maşın "öyrənməsi" arasındakı bu fərq ixtiyari görünsə də - əsasən modellər işlədiyinə görə xoşbəxt olmalıyıq - indi bunun nə üçün belə ciddi təsirlərə malik olduğunu başa düşürük. Mübahisəli şəraitdə süni intellekt sistemləri uğursuzluğa düçar olurlar, "normal" şəraitdə isə çox uğurlu olurlar.

Bunun qarşısını almaq üçün məntiqi addım, modelin öyrəndiyi nümunələrin niyə bu qədər kövrək olduğunu anlamaqdır. Bununla belə, bu, hazırda dərin neyron şəbəkələri kimi ən çox istifadə edilən modellərdə dəstəklənmir, çünki bu modellərin tam olaraq necə və hətta nə öyrəndiyi tam başa düşülmür.

Bu səbəbdən, neyron şəbəkələri kimi süni intellektin əsasını təşkil edən ən məşhur maşın öyrənmə alqoritmləri "qara qutular" adlanır. Biz içəridə nə baş verdiyini bilirik, nəyin çıxarıldığını bilirik, lakin onların arasında nə baş verdiyini bilmirik. Anlamadığımız problemləri etibarlı şəkildə həll edə bilmirik. Eyni şəkildə, bir modelin hücuma məruz qaldığını və ya sadəcə yaxşı olmadığını söyləmək çətin, hətta qeyri-mümkün ola bilər. Qərar ağacları və reqressiya modelləri kimi digər məlumat elmi metodları daha çox şərh və anlayış təmin edə bilsə də, bu üsullar çox vaxt geniş istifadə olunan neyron şəbəkələrinin təmin edə biləcəyi performansı təmin edə bilmir.

Bunu nəzərə alaraq, biz indi bu sistemləri həssas edən AI-nin əsas maşın öyrənməsi alqoritmlərinin xüsusiyyətlərini təyin edə bilərik.

Xüsusiyyət 1: Maşın öyrənməsi yaxşı işləyən, lakin asanlıqla pozulan nisbətən kövrək nümunələri "öyrənməklə" işləyir.

Məşhur inancın əksinə olaraq, maşın öyrənmə modelləri "ağıllı" deyil və tapşırıqlarda, hətta yaxşı yerinə yetirdikləri işlərdə də insan bacarıqlarını həqiqətən təqlid edə bilməz. Bunun əvəzinə, onlar qırılması nisbətən asan olan kövrək statistik assosiasiyaları öyrənməklə işləyirlər. Təcavüzkarlar bu zəiflikdən istifadə edərək əla modellərin performansını məhv edən hücumlar inkişaf etdirə bilərlər.

Xüsusiyyət 2: Verilənlərə tək etibar maşın öyrənmə modellərini pozmaq üçün əsas yol təqdim edir. Maşın öyrənməsi sadəcə verilənlər toplusu adlanan nümunələr toplusundan nümunələr çıxarmaqla “öyrənir”. İnsanlardan fərqli olaraq, maşın öyrənmə modellərinin heç bir əsas biliyi yoxdur - onların bütün bilikləri tamamilə gördükləri məlumatlardan asılıdır. Məlumatların zəhərlənməsi AI sistemlərini zəhərləyir. Bu xarakterli hücum əslində AI sistemini hücumçunun seçdiyi vaxt aktivləşdirə biləcəyi Mançuriya namizədinə çevirir.

Xüsusiyyət 3: Ən müasir alqoritmlərin qara qutu xarakteri auditi çətinləşdirir. Dərin neyron şəbəkələri kimi geniş istifadə olunan ən müasir maşın öyrənmə alqoritmlərinin necə öyrəndiyi və işlədiyi barədə çox az şey məlumdur. Bu gün də onlar bir çox cəhətdən sehrli qara qutu olaraq qalırlar. Bu, maşın öyrənmə modelinin pozulduğunu və ya sadəcə düzgün işləmədiyini müəyyən etməyi çətinləşdirir, hətta qeyri-mümkün edir. Bu xüsusiyyət süni intellekt hücumlarını, hətta tapmaq belə çətin olsa, zəifliklərin yaxşı müəyyən edildiyi ənənəvi kibertəhlükəsizlik problemlərindən fərqləndirir.

Birlikdə götürdükdə, bu zəif cəhətlər süni intellekt hücumlarına qarşı mükəmməl texniki həllin niyə olmadığını izah edir. Bu zəifliklər ənənəvi kibertəhlükəsizlik

pozuntuları kimi düzəldilə bilən “baqlar” deyil. Onlar indiki ən müasir süni intellektin mərkəzində duran problemlərdir.

İndi bu hücumların niyə mümkün olduğunu başa düşdükdən sonra bu hücumların praktiki nümunələrinə baxaq.

Giriş hücumu. Giriş hücumları sistemə daxil olan girişləri dəyişdirərək AI sistemlərinin nasazlığına səbəb olur. Girişə "hücum nümunələri" əlavə etməklə həyata keçirilir, məsələn, kəsişmədə dayanma işarəsini vurmaq və ya sosial şəbəkəyə yüklənmiş rəqəmsal fotoşəkildə kiçik dəyişikliklər etmək buna misal kimi göstərilə bilər.

Giriş hücumları təcavüzkardan AI sisteminə hücum etmək üçün onu güzəştə getməsini tələb etmir. Yüksək dəqiqliyə malik və bütövlüyü, məlumat dəstləri və ya alqoritmləri heç vaxt pozulmamış mütləq ən qabaqcıl süni intellekt sistemləri hələ də giriş hücumlarına qarşı həssasdır. Digər kiberhücumlardan tam fərqli olaraq, hücumun özü həmişə kompüterdən gəlmir!

3.2 Maşın öyrənmə alqoritmləri ilə performans analizinin yoxlanması

Təhlükəsizliyi olmayan bir internet düşünsək, burada hər hansı bir istifadəçi axtarış tariximizi, fəaliyyətimizi, kredit kartı nömrəmizi, hökumət təyinatlı şəxsiyyət vəsiqəmizi, hətta bütün gizlilik məlumatlarımızı izləyə bilərlər. Belə bir halda təhlükəsiz olmayan mühitdən yəni, İnternetdən istifadə edərdik? Çoxlarının buna yox cavabı verəcəyinə əminik. Bu faktorlar katastrofik səslənir, lakin insanlar kompüter şəbəkələrinə qorunmasız şəkildə qoşulduqda bunlar baş verə biləcək şeylərin sadəcə kiçik bir nümunəsidir. Bu günkü dövrdə, istifadəçi gizliliyi, iş sahələri və hökumətlər kritik prioritetlər kimi qəbul edilən informasiya texnologiyalarının əsas sahələrindən biridir. Qədim dövrlərdən bəri məlumatların təhlükəsizliyi hem problem, hem də əsas

məsələ olmuşdur və bu səbəbdən müxtəlif yollar və üsullar tətbiq edilməyə çalışılmışdır, lakin heç biri gizliliyi tamamilə qoruya bilməmişdir. Məsələn, məşhur hərbi komandan və siyasətçi Yulius Sezar, hərbi məktubların şifrəli versiyalarını göndərmək üçün ancaq bir rəqəm olan açarları istifadə edirdi. Açarları rəqəmlər kimi istifadə edərək, sözün bütün hərfləri sol və ya sağa doğru keçirilərək açıqlayıcı bir məna əldə edilirdi. Bu gizlilik seçimləri, insan zəkası inkişaf etdiyi zaman sadə səslənirdi, lakin Yulius Sezarın dövründə düşmənlər bu şifrəli məktubları insanlar üçün anlaşılan dilə tərcümə etməkdə çətinlik çəkirdilər. Dünya Səhiyyə Təşkilatının (DST) rəsmi statistikalarına əsasən, dünya əhalisi illik ortalama 1,05% artır. Dünya əhalisinin sayı çoxaldıqca, onların məlumatlarını idarə etmək, gizliliklərini təmin etmək və təhlükəsiz bir mühit yaratmaq, məlumat və gizlilik həvəskarı olan qara papaqlı hakerlərin məlumatları və şəxsi sənədləri oğurlamasından dolayı daha mürəkkəb hala gəlir.

Data və informasiya iki yolla ələ keçirilir:

- İstifadəçilər aldadılır və istifadəçi adları, parollar və s. ələ keçirilir.
- Qara papaqlı hakerlər müxtəlif alternativ yollarla sistemi hədəfləyir (DDOS, zərərli proqram, s.).

Yuxarıdakı məsələləri həll etmək üçün işçilərə iş məsələləri haqqında təlim keçirilməlidir ki, bu da onların kibertəhlükəsizliyin nə olduğunu, təhlükəsizlik prinsiplərinin necə izlənilməli olduğunu və insanların məsuliyyətsizliyi nəticəsində yaranan məlumat təhlükəsizlik pozuntularının katastrofik nəticələrə gətirib çıxara biləcəyini daha da başa düşülən hala gətirsin. Əks təqdirdə, birinci səbəb aradan qaldırılmayacaq, çünki vətəndaşlar şəxsi məlumat və fayllarını əl ilə üçüncü tərəflərlə bölüşməməkdən doğrudan da məsuldurlar. Müxtəlif yollarla sistemə hücum etmək üçün isə ekspertlər təyin olunmuşdur və qara papaqlıların hücumlarına qarşı sistemə qoruyucu tədbirlər görülməlidir.

Şəbəkə hücumları və onların növləri qeyd fayllarının analiz edilməsi ilə aşkar edilə bilər və sistem aşkar edilən hücumlara əsasən gücləndirilə bilər, beləliklə də bu günün texniki vasitələri və alqoritmləri istifadə edərək edilən intrüziyalar qırıla bilməz .

Kiber təhlükəsizliyə əlavə olaraq, data analitikası və elmə həsr edilmiş başqa bir sahə də maşın öyrənməsidir. "Maşın öyrənməsi, insanların öyrəndikləri və dərəcədə dəyişərək doğruluğunu artıran məlumatlar və alqoritmlərlə məşğul olan süni intellektin və kompüter elminin bir bölməsidir" . Çünki siber hücumlar sistem qeyd fayllarındakı izlərdən aşkar edilə bilər, maşın öyrənməsi alqoritmləri qeyd fayl məlumatlarından bu cür siber hücumlarının nümunələrini öyrənərək, qara şapkaları təkərə salaraq və ya autentifikasiyalarını pozaraq bənzər hücumları qarşılaya bilərlər. Yalnız 2018-ci ildə dünya üzrə 10,5 milyard kötü proqram hücumu müşahidə edilmişdir, insanların idarə etməsi üçün çox böyük bir saydır; Lakin maşın öyrənməsinin köməyi ilə kompüterlər kötü proqram hücumları datasetlərinə təlim alaraq onlardan nümunələri öyrənə və nümunələrdən nümunələr öyrənərək belə hücumları əngəlləməyə kömək edə bilərlər.

Maşın öyrənməsinin perspektivindən də phishinq e-poçtlar asanlıqla aşkar edilə bilər, çünki domain adlarının, linklərin, əlavələrin və düymələrin analizi kompüterlərin onları avtomatik olaraq aşkar etməsini, onları blok etməsini və qeyd fayllarında şəbəkə duvarlarını gücləndirməsini kömək edir. Bu, maşın öyrənməsinin dünya üzrə siber hücumlarla mübarizə aparmağa necə kömək edə biləcəyinin nümunələrindən biridir. Lakin maşın öyrənməsi ilə kömək olaraq bir təhlükəsizlik sistemi həyata keçirmək üçün çox sayda məlumat tələb olunur. Həmçinin, yalnız miqdar problemi həlli edə bilməz, məlumatın keyfiyyəti də vacibdir, çünki maşınlar pis məlumatdan öyrənə bilməz. Bu faktorlara əsaslanaraq, məlumat keyfiyyəti və

miqdarı mühitində cybercrimə və hücumları aşkar etmək üçün təlim modellərinin tədris edilməsinə ehtiyac duyulur və onlara qarşı dərhal və ciddi tədbirlər görülməlidir. Bu məqalə, xüsusiyyətlə sinifləndirmə alqoritmləri olmaqla maşın öyrənmə alqoritmlərinin, iş sahəsində şəxsi məlumatları və gizli sənədləri işləyən serverlərlə əlaqəli sistemlərin cybercrimə və hücumların təsirini azaltmağa necə kömək edə biləcəyini analiz etmək məqsədini daşıyır.

Şəbəkə Təhlükəsizliyində Sinifləndirmə Alqoritmləri. 21-ci əsrin kompüter şəbəkələri bu halı öyrəndikdə hərəkətdə olan qanunsuz fəaliyyəti aşkar etmək çətindir. Lakin, xüsusilə sinifləndirmə alqoritmləri kimi maşın öyrənmə alqoritmləri tərəfindən gücləndirilmiş avtomatlaşdırılmış sistemlər, sistemləri kötü proqram infeksiyalarından və ya zərərçəkən hücumçulardan, spamlardan və phishinq kampaniyalarından qoruyabilərlər. Maşın öyrənməsində bir neçə sinifləndirmə alqoritmi mövcuddur, məsələn, logistiki regresiya, qərar ağacı sinifləndiriciləri, sinir şəbəkələri, XGBoost kimi nümunələr göstərə bilərik. Bu alqoritmlərin effektiv şəkildə işləməsi üçün məlumat keyfiyyəti kifayət qədər yüksək olmalıdır. Daha doğrusu, bu alqoritmlərin siber təhlükəsizlikdə tətbiq edilməsi, parametrlərin hesablanması üçün böyük məlumat miktarının toplanmasını tələb edir. Məlumat toplama və hesablama prosesi şirkətlər üçün saxlama, GPU və Hadoop, Cloudera kimi məlumat ekosistemlərinin mövcudluğu kimi səhvlərə yol açan problemlərlə əlaqəlidir, bu səbəbdən siber təhlükəsizlikdə maşın öyrənmənin tətbiqi digər iş problemlərinə nisbətən geridə qalır, məsələn, maliyyə məsləhətçisi kimi, məhsul tövsiyələri, axtarış optimizasiyası, kimi. Bu faktorlara baxmayaraq, maşın öyrənməsi, xüsusilə də dərin öyrənmə sinifləndirmə alqoritmləri ilə, müdaxilə təyinatı, kötü proqram analizi, spamlar və phishinq kampaniyaları ortasında müdaxilə edə bilər. Müdaxilə təyinatı kompüterlərdə və ya şəbəkələrdə qanunsuz fəaliyyətlərin aşkar edilməsinə işarə edir. IDS-lər müdaxilələrə əsaslanır, lakin hazırda sistemlər anormal hərəkətlərə, təhlükələrə və maşın öyrənmə alqoritmlərinə əsaslanan sinifləndirməyə

kimi əlavə imkanlarla tətbiq edilir. Xüsusilə botnetlərin aşkar edilməsi siniflәndirmə alqoritmləri istifadə edilərək avtomatlaşdırıla bilər. Botnet, əslində, infeksiya olunmuş kompüterlər arasında müşahidə edilən şəbəkə və xarici komanda və nəzarət serverləri arasındakı əlaqələrin aşkar edilməsi üçün hazırlanmışdır. Hərçənd, proseslər avtomatik olabilir, bu məsələlərin vizual analiz üçün əl ilə alətlər istifadə edilməsini tələb edir, bu isə insanlar tərəfindən idarə olunan avtomatik olmayan maşınlar yerinə həyata keçirilir. Beləliklə, maşın öyrənməsi bu problemi həll etmək üçün ideal vasitə ola bilər.

Zərərli proqram analizi, onun haqda toplandığı məlumatları istifadə edərək bir obyektin təhlükə yaradıp yaratmadığını müəyyənləşdirmək üçün istifadə olunur [4].

Zərərli proqram təminatının aşkarlanması adətən iki yolla həyata keçirilir, icradan əvvəl və icradan sonra. Bir neçə onilliklər əvvəl zərərli proqramların identifikasiyası əsasən icradan əvvəl idi. Bununla belə, İnternetdən istifadənin sürətli artımı və zərərli proqram təminatının sonrakı artımı əl ilə aşkarlama qaydalarının artıq mənasını itirdiyi mərhələyə gətirib çıxardı; beləliklə, zərərli proqramların aşkarlanması üçün maşın öyrənmə alqoritmlərindən istifadə edilir [4]. Xüsusiyyət mühəndisliyi üsulları maşın öyrənmə təsnifat alqoritmlərindən istifadə edərək zərərli proqramların aşkarlanmasında ən vacib addımlardan biridir. X xüsusiyyəti fayl məzmununun və ya davranışının bəzi xüsusiyyəti ola bilər [4]. Məsələn, fayl statistikası və istifadə olunan API funksiyalarının siyahısı. Digər tərəfdən, proqnozlaşdırmaq istədiyimiz etiket, başqa sözlə, etiket olaraq Y, 0 və 1 kimi təsnif edilə bilər [4]. Zərərli proqram üçün 1 və qeyri-zərərli proqram üçün 0. Alqoritm və parametrləri seçdikdən sonra siz təlimə başlaya və ən aşağı qiymətə malik olan parametrləri hesablaya bilərsiniz. "Təlim o deməkdir ki, biz xüsusi ölçülərə uyğun olaraq istinad obyektləri toplusunda öyrədilmiş model üçün ən dəqiq cavabları verən xüsusi parametrlər toplusundan istifadə edərək seçilmiş ailədən modeli axtarırıq.

Başqa sözlə, biz tərifi X-dən Y-ə səmərəli xəritəçəkmə üçün optimal parametrlər şəklində öyrənirik" [4].

Spam və fişinq kampaniyalarının aşkarlanması boş vaxt itkisini və e-poçtla bağlı potensial təhlükələri azaltmağa yönəlmiş müxtəlif üsulları əhatə edir [8]. Kibercinayət kimi fişinq zərərli proqramları, yoluxmuş vebsaytlara keçidləri əhatə edə bilər; buna görə də təcavüzkarlar tez-tez bu taktikadan hansısa yolla qarşı tərəfin cihazına giriş əldə etmək üçün istifadə edirlər. Bu faktları nəzərə alsaq, spam və fişinq aşkarlanması getdikcə çətinləşir, çünki təcavüzkarlar filtrlərdən yan keçmək üçün müxtəlif taktikalardan istifadə edirlər.

Bundan əlavə, Bayes təsnifat alqoritmləri e-poçtların spam olub olmadığını müəyyən etməyə kömək edə bilər. Bayes təsnifat alqoritmləri əslində təbii dil emal alqoritmlərindən istifadə edir və söz istifadəsinə əsasən e-poçtun spam olması ehtimalını tapır. Bu gün Outlook və Mail kimi bəzi e-poçt proqramları e-poçtun spam olub-olmadığını müəyyən etmək və proqnozlaşdırılan spamı lazımsız qovluğa yönəltmək üçün təsnifat alqoritmlərini güclü şəkildə tətbiq edir. Bu amillərə əsaslanaraq, təsnifat alqoritmləri kibercinayətkarlıq risklərini minimuma endirmək və kibertəhlükəsizlik üçün daha yaxşı rəqəmsal təhlükəsizlik həlləri təqdim etmək üçün böyük potensiala malikdir.

Nəticədə, müasir kibertəhlükəsizlik təşkilatları sistemin təhlükəsizliyini və onun təhlilini təmin etmək üçün ML alqoritmlərini tətbiq etməyə başlayıblar. Tədqiqatlar göstərir ki, maşın öyrənməsi alqoritmləri kibertəhlükəsizlikdə spesifik davranışların müəyyən edilməsində daha yaxşı nəticələr əldə edir (məsələn, zərərli proqramların təhlili, DDOS hücumları və s.) Potensial qorunma sistemi pozulmalara səbəb olur. Gouveia və başqaları. (2017) qeyd etdi ki, təsnifat üçün bir neçə dərin öyrənmə alqoritmləri, məsələn, B. irəli ötürülən neyron şəbəkəsi, konvolyusiya neyron

şəbəkəsi və təkrarlanan neyron şəbəkəsi təcavüzkarların aşkarlanmasında, xüsusilə də domen yaratma alqoritmlərinin aşkarlanmasında uğurla tətbiq oluna bilər [3] .

Bu amillərə əsaslanaraq, təsnifat alqoritmləri kibercinayətkarlıq risklərini minimuma endirmək və kibertəhlükəsizlik üçün daha yaxşı rəqəmsal təhlükəsizlik həlləri təqdim etmək üçün böyük potensiala malikdir. Nəticədə, müasir kibertəhlükəsizlik təşkilatları sistemin təhlükəsizliyini və onun təhlilini təmin etmək üçün ML alqoritmlərini tətbiq etməyə başlayıblar. Tədqiqatlar göstərir ki, maşın öyrənməsi alqoritmləri kibertəhlükəsizlikdə spesifik davranışların müəyyən edilməsində daha yaxşı nəticələr əldə edir (məsələn, zərərli proqramların təhlili, DDOS hücumları və s.) Potensial qorunma sistemi pozulmalara səbəb olur. Gouveia və başqaları. (2017) qeyd etdi ki, təsnifat üçün bir neçə dərin öyrənmə alqoritmləri, məsələn, B. irəli ötürülən neyron şəbəkəsi, konvolyusiya neyron şəbəkəsi və təkrarlanan neyron şəbəkəsi təcavüzkarların aşkarlanmasında, xüsusilə də domen yaratma alqoritmlərinin aşkarlanmasında uğurla tətbiq oluna bilər [3] .

Şəkil 8. Hücumların ortalama göstəricisi

Hücum növü	F1- nəticəsi	Dəqiqlik
DoS hücumu	0.9955	0.9940
Daşqın (Overflow) hücumu	0.9941	0.9935
SSH kobud güc hücum girişi	0.9918	0.9943
Şübhəli DNS sorğusu	0.9755	0.9955
Gizli giriş zəhərləmə hücumu	0.9678	0.9874
Ümumi yanaşma	0.7987	0.8729

Mənbə. Müəllif tərəfindən tərtib edildi

Bu fərziyyəni yoxlamaq üçün onlar əvvəlki hesabatlardan və Cisco Umbrella-dan məlumat toplayıblar. Spesifik məsələləri müəyyən edərkən göstərici çox yüksək

faiz qaytardı. Eyni şəkildə, çoxmillətli kibertəhlükəsizlik və antivirus təchizatçılarından biri olan Kaspersky nadir və zərərli hücumlarda dərin öyrənmə təsnifatlarından istifadə edir [4]. Daxiletmə funksiyalarının təqdimatlarına əsaslanaraq, onlar eyni vaxtda hər nüsxə üçün birdən çox təsnifatçı hazırlayır, bu da öz növbəsində alqoritmlər nadir hallarda müəyyən növ zərərli proqramları aşkarlayır. Bununla belə, bu spesifik domenlərdə modelin effektivliyini ölçərkən yadda saxlamaq lazımdır ki, yalan pozitiv nisbət aşağı və ya hətta sıfır olmalıdır, çünki bir milyon xoşxassəli faylda bircə yanlış müsbət proqnoz ciddi nəticələrə səbəb ola bilər. bəzi istifadəçilər üçün nəticələr [4].

XGBoost

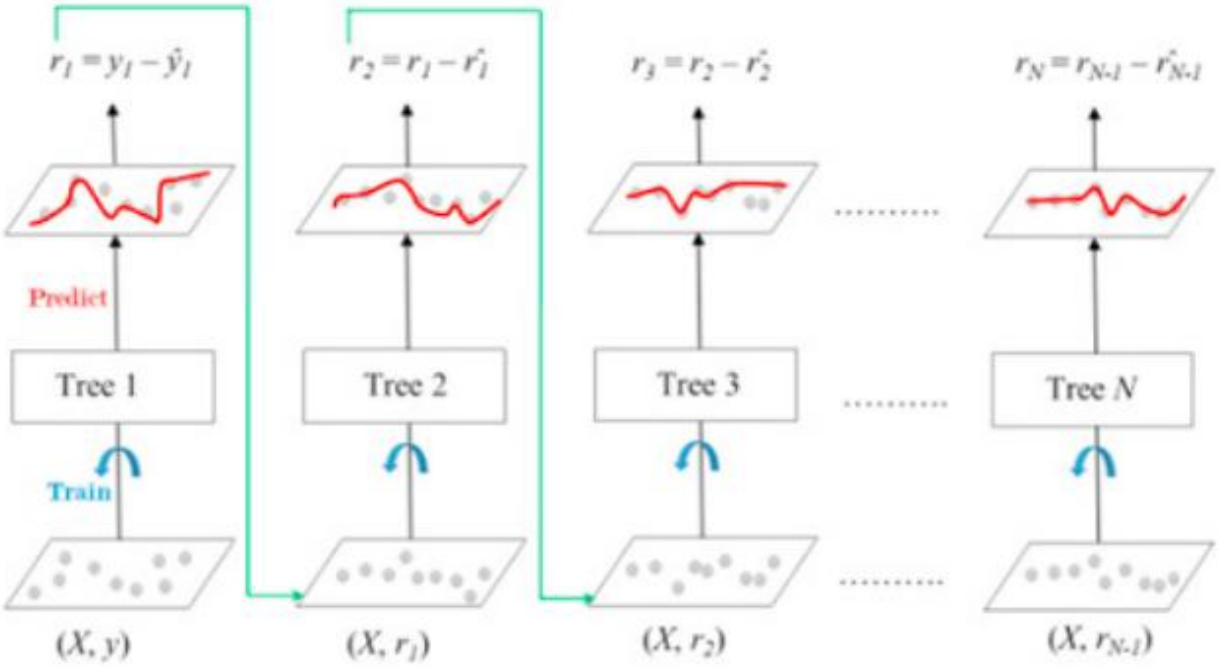
XGBoost ağacın budaqlarına əsaslanan təsnifat üçün ağacdır. Ümumiyyətlə, sadə qərar ağacının özündə bir neçə üstünlükləri var. Aydın biznes vizualizasiyası və şərhi böyük fayda kimi qəbul edilə bilər, çünki rəhbərlər və məhsul sahibləri də biznes layihələrini tamamlayan alqoritmləri başa düşməlidirlər. Qərar ağacları həmçinin davamlı və kateqoriyalı dəyişənləri idarə edə bilər və qaydalara əsaslanan yanaşmadan istifadə etdiyi üçün xüsusiyyət miqyası metodlarını tələb etmir. Bundan əlavə, qərar ağacları çatışmayan dəyərləri avtomatik idarə edə bilər və ümumiyyətlə kənar göstəricilərə davamlıdır. Qərar ağaclarının digər müsbət cəhəti digər alqoritmlərlə müqayisədə məşq etmək üçün daha az vaxt tələb etməsidir. Bununla belə, qərar ağacları həddindən artıq uyğunlaşmaya çox həssasdır [5][8]. Səs-küylü məlumatlar vəziyyətində, yeni qovşaqlar daim yaradılır və şərhi çox mürəkkəbləşdirir.

Bu amildən ötrü yüksək dispersiyanın baş vermə ehtimalı var ki, bu da öz növbəsində son qiymətləndirmədə çoxlu səhvlərə səbəb olur. Bununla belə, modelin yüksək dəyişkənlikdən və ya yüksək qərəzdən əziyyət çəkməsinin qarşısını almaq üçün paketləmə və ya gücləndirici toplama üsulları tətbiq edilir. Artırma ansambl

öyrənməsidir, yəni. H. Hər bir şagirdin performansını fərdi olaraq təkmilləşdirmək üçün çoxlu öyrənənləri cəlb edir [5]. Artırmada n model ardıcıl olaraq öyrədilir, hər bir model əvvəlki modelin uğurunu nəzərə alır, sonrakı modellərdə isə çəkilər artırılır və bu, yüksək xətalara səbəb olur [6].

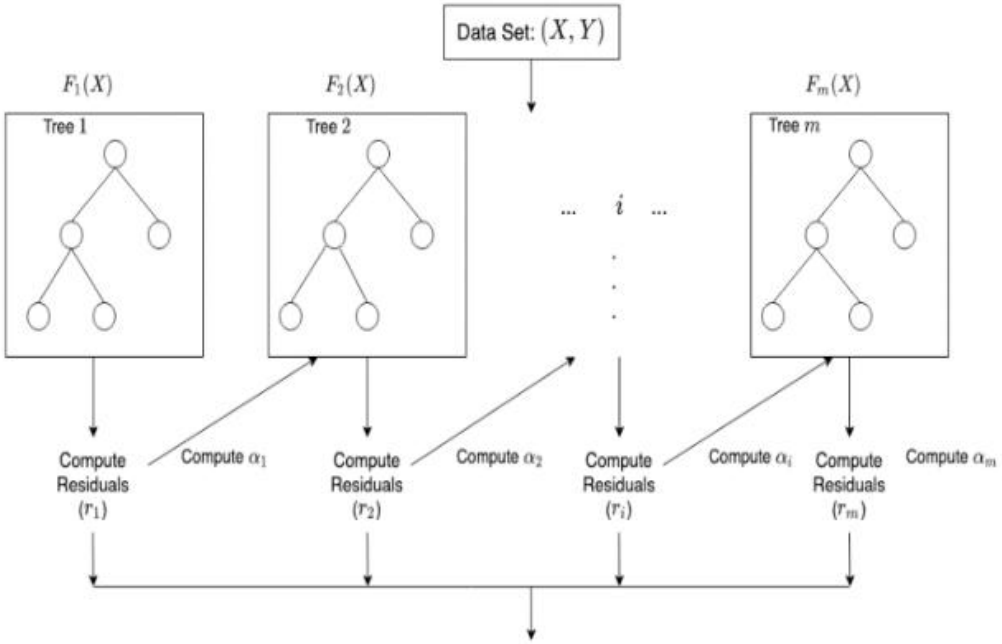
Təlim edilmiş modeldə ən yaxşı performans göstərən fərdi modellər gücləndikcə və son proqnozlara daha çox təsir edir. Təsadüfi meşələrdən fərqli olaraq, gradient gücləndirilməsi yeni proqnozlaşdırıcıları əvvəlki proqnozlaşdırıcıların səhvlərinə uyğunlaşdırmağa çalışır (Şəkil 10). Birgə öyrənmədə gradient qazanc strukturu Şəkil 9-da göstərilmişdir.

Şəkil 9. Birgə öyrənmədə qradient artırma strukturu



Mənbə. G.Abdiyeva- Aliyeva, J.Aliyev, U.Sadigov. Procedia Computer Science 215 (2022) 909–919

Şəkil 10. Birgə (Ensemble) öyrənmədə XGBoost strukturu



Mənbə. G.Abdiyeva- Aliyeva, J.Aliyev, U.Sadigov. Procedia Computer Science 215 (2022) 909–919

Artırma və torbalama arasında bir neçə fərq var. Ağaclar müstəqil şəkildə birləşdirildiyi üçün, əvvəlki modelin səhvlərini öyrənmək və əvvəlki modelin uğursuz olduğu yerdə daha yaxşı performans göstərən yeni model yaratmaq cəhdlərini gücləndirmək. Artırma təhrifi azaltmağa çalışır, torbalama isə həddindən artıq uyğunluğu həll edir. Başqa sözlə desək, "təsnifat qeyri-sabitdirsə, paketləmə tətbiq edilməlidir, lakin təsnifat sabitdirsə və mürəkkəb deyilsə, bu zaman tərsinə artırma tətbiq edilməlidir" [8]. Qradiyent artırma alqoritmi zərər funksiyasının yerli minimumuna çatmaqla xərc funksiyasını minimuma endirmək üçün qradiyent enmə alqoritmindən istifadə edir. Riyaziyyatı sıfırdan gradient gücləndirmədə tətbiq etmək üç əsas giriş tələb edir:

Extreme Gradient Boosting və ya XGBoost yüksək miqyaslana bilən və 2014-cü ildə istifadəyə verilmiş gradient əsaslı qərar ağacı ansamblıdır [5]. Hər bir halda itki funksiyasını minimuma endirməklə əmsalları tapır. Ağacın mürəkkəbliyinə nəzarət etmək üçün itki funksiyasının varyasyonları tez-tez istifadə olunur. XGBoost bu yaxınlarda çox populyarlaşdı və hər iterasiyada daha yüksək sıralı yaxınlaşma istifadə edir. "Extreme Boosting Alqoritminin gücü onun miqyaslı olmasındadır ki, bu da paralel və paylanmış hesablamalar vasitəsilə sürətli öyrənməni təmin edir və yaddaşdan səmərəli istifadəni təmin edir" [7]. Ansambl öyrənmə kimi, təsnifat ağacının ilk əsas modelindəki model sadəcə bir rəqəmi proqnozlaşdırır. Məsələn, bütün girişlərin 0,5 ehtimalı var. İkinci halda, qalıqlar hesablanır. Qalıqları tapdıqdan sonra növbəti mərhələdə oxşarlıq qiyməti aşağıdakı tənlikdən istifadə etməklə hesablanır (tənlik 1).

$$\frac{\sum_1^m (y - \hat{y})^2}{(P(1-P) + \lambda} \quad (1)$$

3.3 Zorla müdaxilənin aşkarlanması sistemləri və qarşısının alınması

Müdaxilənin aşkarlanması sistemlərinin və ya ingiliscə hərfi mənada "Intrusion Detection System" sözlərinin qısaltması olan IDS sistemlərinin əsas mahiyyəti ondan ibarətdir ki, pisniyyətli həmlələri müəyənləşdirsin və qeydə alsın.

"İntrusion Prevention System" sözlərinin qısaltması olan IPS isə azərbaycan dilinə tərcümədə "müdaxilənin aşkarlanaraq qarşısının alınması" mənasını özündə əks etdirir. IPS sistemləri şəbəkə trafikində baş verən pisniyyətli hərəkətləri aşkar edildikdən sonra qarşısını almaq məqsədi ilə icra edilən təhlükəsizlik sistemləridir. Ümumi şəkildə fikrimizi ifadə etmiş olsaq, IDS (müdaxilənin aşkarlanması sistemləri) hücumları aşkarlamağı hədəfləyir, IPS (müdaxilənin qarşısının alınması sistemləri) isə hücumları bloklamaq və qarşısını almaq üçün nəzərdə tutulub.

Son zamanlarda artmaqda olan inkişaf etmiş kiberhücumları aşkarlamaq və qarşısını almaq üçün ciddi şəkildə mahiyyət daşıyır. Bu da öz növbəsində təhlükəsizlik baxımından çox önəmlidir. IDS şəbəkə trafiki zamanı daxil olan paketlərin analizində hücumu aşkar edib loqlayarkən, IPS isə hücumları aşkar etməklə həmin hücumların qarşısını almaq üçün tədbir görür.

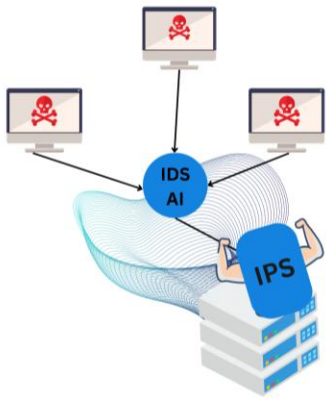
Terminoloji adıyla "Firewall" adlandırılan qoruma divarı şübhəli paketlərin keçidini məhdudlaşdırmış olsa belə hücum ərəfəsində özünü yenidən proqramramlaşdıraraq hücuma hazır etmə funksiyasına malik deyillər. Məhz, elə buna görə də IDS və IPS hər ikisi bir arada olmaqla istifadə edilir ki, təhlükəsizlik tədbirləri daha güclü şəkildə həyata keçsin. Müdaxilənin aşkarlanması sistemlərinin köməyi ilə biz öz korporativ şəbəkəmizə olunan hücumları vaxtında təyin edə və dərhal sonra müdaxilənin qarşısının alınması sistemi vasitəsilə hücumunçunun yenidən sistemimizə müdaxilə etməsinin qarşısını ala bilərik.

Məlumdur ki, biznes proseslərimizin internet texnologiyası ilə şəbəkə üzərindən istifadəsi, bulud texnologiyaları kimi bir çox proqramlardan istifadə bəzi uyğun port və servislərimizi açmaq məcburiyyəti yaradır. Beləliklə, portlarımızı və servislərimizi qorumaq və onların təhlükəsizliyini təmin etmək çətin başa gəlir. Bununla da, şəbəkəmizi tək başına təhlil etməyi çətinləşən qoruma divarı cihazları da zəif reaksiya verməyə başlayır. IPS və IDS sistemləri yeni nəsil Firewall cihazları ilə inteqrasiya olunmuşdur ki, bu da şəbəkə trafikini analiz etmək üçün münbit şərait yaradır.

- Bəs IDS və IPS niyə biryerdə olmalıdır? IDS-in əsas funksiyası loqları tutaraq təhlil etmək olduğu üçün tutduğu loqlara əsasən, hücum növlərini IPS-ə tanıdır. Daha sonra isə IPS həmin loqları cavab tədbiri olaraq bloklayır.

IPS və IDS sistemləri hücumun aşkarlanması və ya təhlilində əsasən iki növ iş məntiqinə malikdir. Birincisi imza əsaslı əməliyyat məntiqi ilə işləyir, ikincisi isə qaydaya əsaslanan əməliyyat məntiqi ilə işləyir.

Şəkil 11. Süni intellektə tanıdılmış hücumun IDS-də detektə olunaraq IPS-lə qarşısının alınması



Mənbə. Müəllif tərəfindən tərtib edildi

Zorla müdaxilə olunmanın aşkarlanması, təhlili və qarşısının alınması sistemləri ümumiyyətlə aşağıdakı xüsusiyyətlərə malikdir;

- Təhlükəsizliklə bağlı administratorlara hücum xəbərdarlığının verilməsi
- Zərərli kodların təyin edilməsi
- Zərərli linklərin təyin edilməsi
- Zərərli paketlərin yoxlanaraq atılması və sıfırlama
- Proqram təminatı və ya istifadəçidən gələn hücumların aşkarlanması
- Müdafiəni gücləndirmək və təkmilləşdirmək üçün hücum nümunələrinin qeydə alınması
- Məhkəmə ekspertizası ekspertləri üçün məhkəmə-tibbi ekspertiza sənədlərinin aparılması
- Məlumatların tam və mövcudluğunun təmin olunması
- Təhlükəsizlik və məxfiliyin təmini

İndi isə zorla müdaxilənin aşkarlanması və qarşısının uğurla alınması üçün süni intellekt metodologiyası ilə birgə nə cür işlənməsindən danışacağıq. Əvvəlcə, bunun nə kimi faydası olmasından danışaq.

Məsələn, 22 portuna "Bruteforce" yəni, kobud güc hücumu göndərilib və biz IDS-ə bu hücumu tanıtmışıq. IDS bu hücumu loqlamağa başlayır və nəticə etibarilə IPS həmin hücumu bloklayır. Həmin hücum 22 deyil də 21 portuna yönələrsə IDS onu loqlamayacaq. Çünki, biz həmin hücumu IDS-ə 22 portu üçün tanıtmışıq. 65535 port növü olduğuna görə bütün portların hamısını tanıtsaq şəbəkədə gecikmələr olacağı nəzərə alınaraq, ən çox rastlaşıla bilinəcək hücumlar süni intellektə tanıdılır. Beləliklə, süni intellekt hücum növlərini tam qavradıqdan sonra süni intellektə qərarvermə səlahiyyəti vermiş oluruq. Növbəti dəfə həmin hücum baş verdikdə süni intellekt həmin hücumu tanıyır. Bununla da, IPS-ə həmin hücumu bloklama əmri verəcək və IPS bu hücumun bloklanması əmrini icra edəcək. Korporativ mövcud olan serverə mütəmadi olaraq gələcək sorğuları süni intellektə öyrətdikdə süni intellekt anlayır ki, hansı sorğular normaldır, hansı sorğular isə təhdid əsaslıdır. Normalda hər hansısa istifadəçi giriş etdiyi zaman bu giriş bir istifadəçidə maksimum 1 dəqiqə və ya minimum 50 saniyə çəkirdisə , növbəti dəfə bu 1 saniyədən də az vaxt alırsa süni intellekt bunu artıq anomaliya kimi qiymətləndirir. Nəticədə süni intellekt IPS-ə bloklama əmri verir.

İki istifadəçi bir-biri ilə paket marşrutlaşdırılması həyata keçirdiyi zaman həmin sorğular müəyyən edilmiş alqoritm əsasında reallaşır. Yəni, müəyyən limitlər təyin edilir. Həmin limitlərə uyğunsuzluq halı baş verdikdə bu artıq anomaliya kimi müəyyən edilir. Biz bilirik ki, hücum anında insan əli ilə oluna biləcək hərəkətlərin avtomatlaşdırılmış şəkildə reallaşdırılması 1 saniyədən tez vaxtda ola bilər. Bu cür hallar təyin edilmiş alqoritmlə sorğulandıqda vaxtında və operativ aşkarlama və qarşıdurma baş verə bilər. Bunun üçün konfiqurasiya fayllarında hücumlar qeyd

edilməlidir və hansı zamanlama anında olduqda normal, hansı zaman aşımını keçdikdə isə anormal olduğu ilə bağlı qeydlər edilməlidir.

Hücumçu hücum obyektini hədəfə aldıqda bir çox yollara əl ata bilir. Buna görə də qabaqlayıcı tədbirlər görmək üçün əvvəlcədən sistemimizin ən çox qarşılaşa biləcəyi hücumları ehtimal edərək onları qeydə almalıyıq.

"Cybernetics" şirkətinin mail domeni vasitəsilə bir nümunə göstərərək real mail domeninə olunan hücumu göstərə bilərik ki, bu da proseslərin nə cür həyata keçdiyini nəzərdən keçirmək üçün şəkil 12-də öz əksini tapmışdır. Şəkilə baxaraq proseslərin nə cür reallaşmasının şahidi ola bilərik.

Şəkil 12. IDS və IPS vasitəsilə hücumların qarşısının alınmasına dair mümunə.

Mənbə. Müəllif tərəfindən tərtib edildi

Şəkil 12.-də 7 may 2023-cü ilin giriş hesabatına fikir versək, 84.54.50.197 ip ünvanına malik girişi loqlayaraq 3 cəhdlə həyata keçirilən uğursuz girişin 1 saniyədən tez müddətdə icra edildiyini görüb, şübhəli olduğunu müəyyən edən IDS bununla bağlı IPS sisteminə siqnal ötürür və beləliklə, IPS həmin ünvanı 600 saniyəlik blok edir. Növbəti dəfə 84.54.50.199 ip ünvanı ilə giriş etməyə cəhd edilir. Bu dəfə də şübhəli əməliyyatın 1 saniyədən az olduğunu görən IDS yenidən IPS-ə

şübhəli girişlə bağlı məlumat ötürür. IPS isə bu ip ünvanı ilə edilən girişi də 600 saniyəlik blok edir.

Şəkil 13. Süni intellektə tanıdılmış hücumların IDS və IPS vasitəsilə qarşısının alınması

05/01/2023, 06:11:01 PM	crit	Banning 141.98.11.65/32 for 240 minutes
05/01/2023, 06:11:01 PM	warn	141.98.11.65 matched rule id 3 (warning: unknown[141.98.11.65]: SASL LOGIN authentication failed: UGFzc3dvcmQ6)
05/01/2023, 06:06:16 PM	crit	Banning 91.224.92.110/32 for 240 minutes
05/01/2023, 06:06:16 PM	warn	91.224.92.110 matched rule id 3 (warning: unknown[91.224.92.110]: SASL LOGIN authentication failed: UGFzc3dvcmQ6)
05/01/2023, 05:48:12 PM	crit	Banning 141.98.11.52/32 for 240 minutes
05/01/2023, 05:48:12 PM	warn	141.98.11.52 matched rule id 3 (warning: unknown[141.98.11.52]: SASL LOGIN authentication failed: UGFzc3dvcmQ6)
05/01/2023, 05:47:46 PM	crit	Banning 141.98.11.29/32 for 240 minutes
05/01/2023, 05:47:46 PM	warn	141.98.11.29 matched rule id 3 (warning: unknown[141.98.11.29]: SASL LOGIN authentication failed: UGFzc3dvcmQ6)
05/01/2023, 05:41:34 PM	crit	Banning 45.125.65.37/32 for 240 minutes
05/01/2023, 05:41:34 PM	warn	45.125.65.37 matched rule id 3 (warning: unknown[45.125.65.37]: SASL LOGIN authentication failed: UGFzc3dvcmQ6)
05/01/2023, 05:38:04 PM	warn	4 more attempts in the next 600 seconds until 141.98.11.111/32 is banned
05/01/2023, 05:38:04 PM	warn	141.98.11.111 matched rule id 3 (warning: unknown[141.98.11.111]: SASL LOGIN authentication failed: UGFzc3dvcmQ6)

Mənbə. Müəllif tərəfindən tərtib edildi

NƏTİCƏ VƏ TƏKLİFLƏR

Texnoloji inkişaf davam etdikcə hücumlar inkişaf edərək daha təkmilləşmiş formada özünü biruzə verməyə başlayır. Bu da təhlükəsizlik baxımından böyük narahatlıq doğurur. Kiber təhdidlərə qarşı yüzdə-yüz qabaqcıl tədbir almaq mümkün deyildir. Lakin, bu təhdidlərin gətirə biləcəyi ziyanlara qarşı daha az zərərlə vəziyyəti ələ almaq mümkündür. Tədqiqat işində aparılan təcrübələrə dayanaraq deyə bilərik ki, məxfi məlumatlarımız qarşı tərəf üçün maraqlı olduqda, kiber hücumçu bunu bir çox variantlara əl atmaqla həyata keçirə bilər. Bu səbəbdən də, həmin hücum anını qabaqcadan düşünmək və hücum anını gözləmədən buna qarşı bir çox atıla biləcək qabaqlayıcı tədbirlərə fokslanmaq lazımdır. Bunları nəzərə alaraq günümüzdə istifadə olunan insident aşkarlama və müdaxilə sistemlərinin tətbiqində süni intellekti tətbiq edərək yeni hazılanmış zərərli kodlara qarşı dayanıqlı olmaq zəruridir. Süni intellektin tətbiqi kibertəhlükəsizlik həllərində insan faktorunu azaldaraq kiberhücumlara qarşı daha çevik reaksiya verməyi və qabaqlayıcı tətbirlər görməyi təmin edir. Günümüzdə istifadə olunan IDS, IPS, SIEM, EDR, XDR kimi həllərdə 10 minlərlə təhlükəsizlik siyasəti tətbiq olunurki bu siyasətlərin yeni nəsil versiyalarının hazırlanması və aktuallığının analiz olunması ilə müasir təhdidlərə qarşı dayanıqlı olması üçün süni intellektin təhlükəsizlik sistemlərində tətbiqi zəruridir.

Təhlükəsizlik sistemlərində istifadə olunanan siyasətlərin alqoritmlərinin fəaliyyəti laboratoriya mühitində analiz olunaraq maşın öyrənmə texnologiyası ilə öyrədilməli və yeni siyasətlərin hazırlanmasında süni intellektin tətbiqi üzrə təhlükəsizlik sistemləri inkişaf etdirilməlidir.

İXTİSARLAR

AI - Artifical Intelligence (Süni İntellekt)

DoS - Denial of Service (Xidmətdən İmtina)

DDoS - Distributed Denial of Service (Paylanmış Xidmətdən İmtina)

IP - Internet Protocol (İnternet Protokolu)

IDS - Intrusion Detection (System Müdaxilənin Aşkarlanması Sistemi)

IPS - Intrusion Prevention System (Müdaxilənin Qarşısının alınması Sistemi)

SIEM -Security İnformation and Event Managment (Təhlükəsizlik Məlumatı və Hadisələrin İdarə Edilməsi)

Sİ- Süni intellekt

EDR - Endpoint Detection and Response (Son nöqtə aşkarlama və cavab)

XDR - Exdended Detection and Response (Genişləndirilmiş aşkarlama və cavab)

MiTM - Meet-in-The-Middle (Ortada görüş hücumu)

İSTİFADƏ EDİLMİŞ ƏDƏBİYYAT SİYAHISI

1) Sosial mühendislik atakları ve alınması gereken önlemler http://indexive.com/uploads/papers/pap_indexive15949787772147483647.pdf

2) https://web.itu.edu.tr/~sonmez/lisans/es/uzman_sistemler_giris.pdf

3) Htun, P. T. ve Khaing, K. T. (2013). Detection model for daniel-of-service attacks using random forest and k-nearest neighbors. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 2(5), 1855-1860.

4) Özgür, A. ve Erdem, H. (2018). Saldırı tespit sistemlerinde genetik algoritma kullanarak nitelik seçimi ve çoklu sınıflandırıcı füzyonu. Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi, 33(1).

5) Sasikala, D., and K. Venkatesh Sharma. "Deployment of Artificial Intelligence with Bootstrapped Meta-Learning in Cyber Security." Journal of Trends in Computer Science and Smart Technology 4, no. 3 (2022): 139-152.

6) Karunakaran, P. "Deep Learning Approach to DGA Classification for Effective Cyber Security." Journal of Ubiquitous Computing and Communication Technologies (UCCT) 2, no. 04 (2020): 203-213.

7) G. Abruzzese, M. Colajannim L. Feretti, On the effectiveness of machine and deep learning for cyber security, 10th International Conference on Cyber Conflict (CyCon), 2018. DOI: 10.23919/CYCON.2018.8405026

Dolayı istinadlar

[1]https://www.ntv.com.tr/galeri/teknoloji/korkutan-teknoloji-deepfake-ne-demek,5_D8orRnW0KXZgFvaz6Jmw/QKd9LUPuzUmZrX4Be6IocQ

[2]https://www.warhistoryonline.com/world-war-ii/fusag-theghost-army-pattons-d-day-force-that-was-only-a-threat-xb.html.

[3] A. Gouveia, M. Correia, Network Intrusion Detection with XGBoost, IBM Belgium, 2017. https://www.gsd.inescid.pt/~mpc/pubs/XGBoost_chapter.pdf

[4] Kaspersky, Machine Learning for Malware Detection. Official Documentation of Kaspersky. 2021.https://media.kaspersky.com/en/enterprisesecurity/Kaspersky-Lab-Whitepaper-Machine-Learning.pdf

[5] C. Bentejac, A. Csörgo, G. Martinez-Munoz, A Comparative Analysis of XGBoost, 2019., pp.3-10. https://arxiv.org/pdf/1911.01914.pdf

AWS, Amazon documentation. How XGBoost Works, 2018. https://docs.aws.amazon.com/sagemaker/latest/dg/xgboost-HowItWorks.html

[6] M. Hajaram N.F. Jumaat, M.A.Bakar, Boosting the Accuracy of Phishing Detection with Less Features Using XGBoost, International Journal of Software & Hardware Research in Engineering, Vol. 8 (2), 2020. https://www.researchgate.net/publication/339727375

[7] T. Chen, C. Guestrin, XGBoost: A Scalable Tree Boosting System, University of Washington, 2016. https://www.kdd.org/kdd2016/papers/files/rfp0697-chenAemb.pdf

[8] M.Ahasan, R.Gomes, M.M.Chowdhury, K.E.Nygard, Enhancing Machine Learning Prediction in Cybersecurity Using Dynamic Feature Selector, Journal of Cybersecurity and Privacy, 2021, https://doi.org/10.3390/jcp1010011

[9] G. Abruzzese, M. Colajannim L. Feretti, On the effectiveness of machine and deep learning for cyber security, 10th International Conference on Cyber Conflict (CyCon), 2018. DOI: 10.23919/CYCON.2018.8405026

[10] https://en.wikipedia.org/wiki/Exploit_(computer_security)

[11] https://terranovasecurity.com/examples-vishing/

[12] https://www.siberay.com/sosyal-muhendislik

İSTİFADƏ OLUNMUŞ ŞƏKİLLƏRİN SİYAHISI