Süni İntellekt Əsaslı Dezinformasiyanın Yaradılması, Tətbiqi və Onunla Mübarizə

Ziyafət Əmirov
Dövlət Təhlükəsizliyi Xidmətinin Heydər Əliyev adına Akademiyası
zamirov@gmail.com

***Xülasə*— Müasir dövrün qlobal informasiya müharibələri kontekstində saxta məlumatların və xəbərlərin yayılması ictimai şüurun manipulyasiyası üçün əsas vasitələrdən birinə çevrilmişdir. Saxta məlumatlar milli təhlükəsizliyə və kibersuverenliyə təhdid yaradır, dövlət qurumlarına etimadı sarsıdır və sosial mühitdə çeşidli problemlərin yaranmasına səbəb olur. Süni intellektdən istifadə ilə saxta məlumatların yaradılması, istifadəsi və onlara qarşı effektiv mübarizənin aparılması məsələsi aktuallaşır. Məqalədə saxta xəbər, dərin saxta (deepfake), süni intellekt agenti anlayışları, habelə süni intellekt əsaslı dezinformasiyanın yaradılması mexanizmləri, onun kibertəhlükəsizlik mühitində tətbiqi və yaratdığı risklər təhlil edilir. Dezinformasiyanın aşkarlanması üçün istifadə olunan texniki və təşkilati müdafiə yanaşmaları, media savadlılığı və hüquqi tənzimləmələrin rolu şərh edilir.**

Açar sözlər*— *milli təhlükəsizlik; rəqəmsal suverenlik; kibersuverenlik; dezinformasiya; saxta xəbər; dərin saxta (dipfeyk); süni intellekt agenti.

I. GİRİŞ

Müasir informasiya cəmiyyətində süni intellekt (Sİ) texnologiyalarının sürətli inkişafı informasiya istehsalı və yayılmasında köklü dəyişikliklərə səbəb olmuşdur. Bu dəyişikliklər bir tərəfdən məhsuldarlığı və əlçatanlığı artırsa da, digər tərəfdən dezinformasiyanın daha sürətli, real və kütləvi şəkildə yaradılması üçün yeni imkanlar formalaşdırmışdır. Sİ əsaslı mətn, şəkil, səs və video generasiya alətləri (deepfake texnologiyaları daxil olmaqla) ictimai rəyə təsir, siyasi manipulyasiya, kiberhücumların gücləndirilməsi və sosial etimadın sarsıdılması kimi ciddi risklər yaradır. Bu baxımdan, süni intellekt əsaslı dezinformasiyanın yaradılma mexanizmlərinin, tətbiq sahələrinin və onunla mübarizə üsullarının elmi şəkildə araşdırılması kibertəhlükəsizlik və informasiya təhlükəsizliyi üçün xüsusi aktuallıq kəsb edir.

II. SAXTA XƏBƏRLƏR VƏ ONLARA NÜMUNƏLƏR

"Saxta xəbər" anlayışının sabit bir tərifi yoxdur. Xəbər kimi təqdim edilən yalan məlumata həmçinin, yüksək vəzifəli insanlar tərəfindən özlərinə xoş olmayan hər hansı bir xəbərə tətbiq etmək üçün istifadə edilmişdir. Dezinformasiya isə zərərli niyyətlə saxta xəbər yaymağı əhatə edir, xüsusən də seçkilər zamanı düşmən xarici aktorlar tərəfindən yaradılır və təbliğ olunur. Bu anlayışın mənası belədir: "saxta xəbərlər" tamamilə uydurma, təsdiqlənə bilən faktlar, mənbələr və ya sitatlar olmayan xəbərlərdir.

Bu anlayış siyasiləşib və hər hansı bir əks baxış bucağını nüfuzdan salmaq üçün geniş istifadə olunur. Bəzi insanlar saxta xəbərlərdən rəqiblərinə, mübahisəli məsələlərə və ya media təşkilatlarının etibarlılığına şübhə ilə yanaşmaq üçün istifadə edirlər. Sosial media, botlar və süni intellektdən (Sİ) geniş istifadə kimi texnoloji irəliləyişlər saxta xəbərlərin tez və asanlıqla yayılmasına imkan verir.

Saxta xəbərlər əsasən iki qrupa ayrılır [1]:

- Yanlış məlumat (misinformation) - səhv nəticəsində və ya təsadüfən yaradılmış və yayılmış yalan və ya qeyri-dəqiq məlumatdır. Onun məqsədi aldatmaq deyildir.
- Dezinformasiya (disinformation) - dezinformasiya ictimai rəyə təsir göstərmək və ya həqiqəti gizlətmək üçün qəsdən yaradılan və yayılan yalan məlumatdır.

Yanlış məlumat və dezinformasiyanın yaranma səbəblərinə aşağıdakı nümunələri göstərmək olar:

- Seçki kampaniyalarında siyasi mənfəət üçün seçicilərə və siyasətçilərə təsir göstərmək və ya ictimai müzakirələrə təsir etmək istəyi. Məsələn, seçki saxtakarlığı haqqında qəsdən dezinformasiyanın yayılması;
- Daha çox klik ilə daha çox pul əldə etmək məqsədi. Bəzi xəbər hekayələri, məzmunundan asılı olmayaraq, maliyyə mənfəəti üçün klik yaratmaq istəyən insanlar tərəfindən yaradılır. Klik fırıldaqçılığında istifadə olunan alət və üsullara aşağıdakı nümunələri göstərmək olar: klik-fermaları (click farms); botnetlər (botnets); klikləmə (clickjacking); klik inyeksiyaları (click injection); reklam yığılması (ad stacking).
- Siyasi rejimlərin öz təbliğatlarını inkişaf etdirmək istəyi (məsələn, Rusiyanın Ukrayna ərazilərini işğalı ətrafındakı "Xüsusi hərbi əməliyyat" hekayəsini idarə etmək üçün "saxta xəbərlər"i yayması).

Saxta paylaşımlara gəldikdə isə, onlar çox vaxt faktlara əsaslanır, lakin bəzi məlumatları təhrif edir və ya gizlədir, reallığın yanlış mənzərəsini yaradır. Sosial şəbəkələrdə belə paylaşımlara maraqlı nümunələr göstərmək olar [2]:

1) Facebook-da saxtakarlıq №1. Yaşlı bir alman qadını gənc qaçqınla evlənir.
Facebook-da Avropa haqqında bir neçə yüksək profilli saxtakarlıqlar peyda olmuşdur. Məsələn, "Narodny+" adlı bir "əyləncə bloqu" 3913 qarşılıqlı əlaqə toplayaraq oxucuları Almaniyada 85 yaşlı bir qadının 18 yaşlı qaçqınla evlənməsi xəbəri ilə heyrətləndirmişdi;

DOI: 10.25045/cyber.gov.2026.28

2) Twitter-də (X-də) saxtakarlıq №3. Almaniya Rusiyadakından da dəhşətli pensiya islahatı keçirir.
Almaniya, Rusiya ilə birlikdə, pensiya yaşını 67-yə qaldırdı. Bunun üçün 1517 qarşılıqlı əlaqə yaratmış blogger Kisliçenko iddia edir ki, tezliklə Almaniyanın pensiyaları azaldılacaq və pensiya yaşı artacaq. Əslində bu iddia ilə bağlı hər şey yalandır. Birincisi, Almaniyada pensiya yaşının artırılması qərarı 2007-ci ildə verilib və 2012-ci ildə qüvvəyə minib. Almaniyada ömür uzunluğu və onun Rusiya ilə müqayisəsi barədə danışmaq isə yersizdir - hər şey aydındır.

Hər gün kiber məkanda iqtisadi, siyasi, hərbi və digər sahələri əhatə edən çoxsaylı saxta məlumatlarla rastlaşırıq. Dünyanın diqqət mərkəzində olan Rusiya-Ukrayna müharibəsinə dair bir neçə nümunəni təqdim etməklə kifayətlənəcəyik. Stopfake.org komandası Rusiya təbliğatının ən təəccüblü saxtakarlıqlarını təqdim etmişdir [3]:

1) Zoryan və Şkiryak. Ən gülməli saxtakarlıq, Ukrayna siyasətçisi və Ukrayna Daxili İşlər Nazirliyi rəhbərinin müşaviri Zoryan Şkiryakın populyarlıq qazandırdığı yalan idi. Rusiya mediası onu Donetsk döyüşçüsü Givinin qətlində günahkar olduğu iddia edilən iki nəfərdən biri kimi təqdim etmişdir. Məsələn, "Rossiya-24" kanalı Zoryan və Şkiryakın Givinin qətlini təşkil etdiyini və cinayətkarın "hansısa bir Yaroş" olduğunu bildirmişdi;

2) Ukraynada ac insanlar göyərçinlərdən çörək alırlar. "Ukraynada 2017-ci ilin aclıq qışı" mövzusu Rusiya təbliğatı üçün əsas mövzu olur. Bu istiqamətdə ən çox müzakirə Rusiya Müdafiə Nazirliyinin Zvezda televiziya kanalında "Ukraynada ac insanlar göyərçinlərdən çörək alırlar" mövzusunda dərc olunmuş məqalə olur;

3) Ukrayna - ABŞ-ın bioloji sınaq poliqonu. Bir sıra Rusiya və xarici media orqanları Ukraynada gizli ABŞ bioloji təcrübələri haqqında məlumat yayıb və onları eyni media orqanlarında müzakirə olunan müxtəlif xəstəliklərin baş verməsi ilə əlaqələndiriblər. Bəzi Rusiya mediası uşaqlar arasında xəstəlik baş verməsi kontekstində "Pentaqonun Avropaya qarşı bioloji bombaları" haqqında yazan və UNICEF-in Giovanna Barbarisdən sitat gətirən Global Research-dən "Qərb jurnalistləri"nə (əslində tanınmış təbliğatçılar) istinad edib. Lakin, UNICEF nümayəndəsi öz hesabatında xəstəliyin baş verməsinin başqa bir səbəbinə işarə edir: uşaqlar arasında peyvənd nisbətinin aşağı olması.

Saxta məlumatların yayılmasında Belarus Rusiyadan heç də geri qalmır. Belə paylaşımlardan birini qeyd edirik [4]: Polşa ordusunu modernləşdirmək üçün 540 milyard dollar kredit götürüb. Polşa 2023-cü ilin oktyabr parlament seçkiləri ərəfəsində Belarusun yaratdığı təhlükəni bilərəkdən şişirtməklə yanaşı, daha çox silah almaq üçün böyük məbləğdə kreditlər də götürür. Bu açıqlama 30 avqustda Belarus 1 telekanalının "Trends" proqramında verilib. Açıqlamada bildirilir: "Varşava kreditlə Avropanın ən güclü ordusunu qurur. Onlar ən son hərbi texnika növlərini aktiv şəkildə alırlar: təyyarələr, sualtı qayıqlar, tanklar və s. Polşa hökuməti Cənubi Koreya bankından 200 milyard dollar borc alıb. Və bu yaxınlarda daha 340 milyard dollar istəyib. Bu vəsaiti kimin geri qaytaracağı bəlli deyil". Əslində, Polşa Prezidenti Andjey Duda 14 iyulda Polşanın Cənubi Koreya istehsalçılarından təxminən 9 milyard dollar dəyərində silah sifariş etdiyini açıqladı. Polşa da müqavilənin maliyyələşdirilməsi üçün Cənubi Koreyadan eyni məbləğdə kredit tələb etmişdi. Bu rəqəmlər Koreya televiziyası tərəfindən də təsdiqlənmişdir.

III. DƏRİN SAXTA/DİPFEYK VƏ ONA NÜMUNƏLƏR

Rəqəmsal texnologiyaların sürətli inkişafı informasiya mühitində yeni imkanlarla yanaşı, ciddi risklər də yaradıb. Bu risklərdən biri də son illərdə geniş yayılmağa başlayan deepfake məlumatlardır. Deepfake texnologiyası reallıqla saxtanı bir-birindən ayırmağı çətinləşdirərək cəmiyyət, media və fərdlər üçün təhlükəli nəticələr doğura bilər və ölkə olaraq biz bü cür təhlükəli məqamlarla tez-tez qarşılaşırıq.

Deepfake, media tərəfindən təsvir edilən bir şəxsin başqası kimi görünməsi üçün süni intellekt tərəfindən yaradılan və ya manipulyasiya edilən bir media parçasıdır. Bu, bir görüntünün, audio parçanın, videonun və ya bunların hər hansı bir kombinasiyasının manipulyasiyasını əhatə edə bilər. "Deepfake" "dərin öyrənmə (deep learning)" və "saxta (fake)"nın qarışığıdır [5].

Şəkilləri, səsləri və ya videoları redaktə etmək üçün vasitələr yeni deyil. Lakin deepfake texnologiyası maşın öyrənmə üsullarından istifadə edərək medianı o qədər dəqiq şəkildə düzəldə və ya dəyişdirə bilər ki, onu qanuni bir şeydən ayırmaq olduqca çətin olar.

Dövlət başçısı və məmurlar adından yayılmış saxta məlumatlar birbaşa dövlətimizə qarşı təhdid kimi qiymətləndirilir və onlara qarşı müvafiq tədbirlər görülür. Nümunə olaraq Azərbaycan Respublikasının Prezidenti Zati-aliləri cənab İlham Əliyevin adından müxtəlif illərdə sosial şəbəkələrdə və messencerlərdə yayılmış çoxsaylı süni intellekt əsaslı dezinformasiyalardan birini qeyd edək: "Səfərbərlik elan olunur" adlı deepfake çıxışı 2023-cü ilin sentyabrında Ermənistanla sərhəd gərginliyi fonunda yaradılmış, TikTok, WhatsApp və Telegram kanallarında yayılmışdır.

Çıxışda iddia olunur ki, ümumi səfərbərlik elan edilir, 18–55 yaş arası kişilər hərbi komissarlığa çağırılır və sərhədlər bağlanır. Video dövlət qurumları tərəfindən saxta elan edildi, informasiya-psixoloji əməliyyat kimi qiymətləndirdi və müvafiq təkziblər verildi.

Türkiyə Respublikasının Prezidenti cənab Rəcəb Tayyib Ərdoğanın adından da çoxsaylı dezinformasiyalar paylaşılmışdır. Nümunə olaraq "Ankara erməni cəmiyyətinin həssas təbəqələrini dəstəkləmək üçün 10 milyon dollar istiqamətləndirəcək" adlı paylaşımı göstərmək olar [6].

Orada yazılır: Türkiyə mediası xəbər verir ki, Ankara "həlak olan hərbçilərin ailələri, təqaüdçülər və xüsusi qayğıya ehtiyacı olan ailələr də daxil olmaqla, Ermənistan cəmiyyətinin həssas təbəqələrini dəstəkləmək üçün" 10 milyon dollar ayıracaq.

Türkiyə Prezident Administrasiyasının Dezinformasiyaya Qarşı Mübarizə Mərkəzi bu kimi paylaşımlara vaxtında reaksiya verərək bəzi media orqanlarında və sosial mediada Türkiyə Prezidenti Rəcəb Tayyib Ərdoğanın həlak olan erməni hərbçilərinin ailələrinə və ehtiyacı olan ermənilərə 10 milyon

dollar ayırmaq barədə fərman imzaladığı barədə yayılan məlumatları bəyanatla təkzib etmişdir [7].

Hazırkı qabaqcıl deepfake süni intellekt bir-birinə qarşı işləyən iki maşın öyrənmə modeli ilə təchiz edilmişdir. "Generator" alqoritmi nümunə görüntüləri, audio və videodan istifadə edərək nümunələrə mümkün qədər bənzəyən yeni bir media parçası yaratmaq və ya mövcud olanı manipulyasiya etmək üçün öyrədilir.

"Ayrı-seçkilik (discriminator)" alqoritmi nümunələrdəki fərqli xüsusiyyətləri tanımaq və "generator"un onları harada qaçırdığını göstərmək üçün təlim keçib ki, geri qayıdıb bu uyğunsuzluqları düzəldə bilsin. Bu proses belə gedir [5]:

1) *Generator və diskriminator alqoritmləri media nümunələrindən məlumatları təhlil edir;*
2) *Generator nümunələri mümkün qədər bənzətmək üçün media yaradır (və ya manipulyasiya edir). Bu, ilkin deepfake-dir;*
3) *Diskriminator nümunələr və deepfake arasında uyğunsuzluqları yoxlayır;*
4) *Generator diskriminatorun deepfake-də tapdığı uyğunsuzluqları düzəldir və deepfake-i diskriminatora yenidən təqdim edir;*
5) *Diskriminator artıq uyğunsuzluq tapana qədər 3 və 4-cü addımları təkrarlayır.*

Bu proses generatora nəticədə medianı o qədər dəqiq yaratmağa və ya manipulyasiya etməyə imkan verir ki, nə süni intellekt, nə də insan zəkası dərin saxta ilə onun əsaslandığı orijinal media arasındakı fərqi asanlıqla ayırd edə bilmir.

Süni intellekt modellərinin hazırlanması üçün lazım olan çoxlu sayda verilənlər çox vaxt ictimaiyyətə açıq mənbələrdən toplanır. Kibercinayətkarlar adətən aşağıdakıları araşdırırlar [8]:

- Korporativ veb saytları (Corporate Websites): işçilərin şəkilləri və tərcümeyi-halları, tanıtdırıcı videoları, qeydə alınmış vebinarlar və rəhbər müsahibələri olan "Haqqımızda" səhifələrinin təqdim etdiyi yüksək keyfiyyətli mənbələri;
- Sosial media (Social Media): LinkedIn, Facebook, Instagram, X, TikTok və YouTube və s. social şəbəkələrdəki profillərdən şəkilləri, videoları və audio yazıları;
- İctimai çıxışlar (Public Appearances): çıxışlar, konfrans təqdimatları, xəbər müsahibələri və onlayn mövcud olan podkastlar (audio və ya video formatında epizodlar şəklində yayımlanan rəqəmsal yayım);

Daha çox yayılmış deepfake texnikalarının növləri bunlardır**:**

- Üz dəyişdirmə (Face Swapping): videoda bir şəxsin üzünün digərinin üzü ilə əvəz edilməsi;
- Səs klonlaşdırma (Voice Cloning): əvvəlcədən yazılmış videoda və ya canlı söhbət zamanı bir şəxsin səsinin sintez edilərək ona istənilən şeyin dedizdirilməsi;
- Dodaq sinxronlaşdırması (Lip-Syncing): mövcud videonu dəyişdirərək, şəxsin əslində heç vaxt demədiyi bir şeyi dediyi kimi göstərilməsi;
- Bütün bədənin yenidən tərtib edilməsi (Full-Body Reenactment): saxta jestlər yaratmaq üçün bir şəxsin bütün bədən duruşunun və hərəkətinin manipulyasiya edilməsi.

Deepfake texnologiyası ilə saxta məlumatların hazırlanması

Deepfake texnologiyasından istifadə ilə süni intellekt əsaslı dezinformasiyaların yaradılması elə də asan məsələ deyildir. Bu prosesdə istifadə olunan alət və vasitələrə aşağıdakı nümunələri göstərmək olar:

- DeepFaceLab - Hollivud səviyyəsində üz dəyişdirmə (face-swap);
- HeyGen/Synthesia - şəkil və ya videonun səsini dəyişərək onu fərqli dillərdə səsləndirmək;
- ElevenLabs - səs klonlama. Cəmi 30 saniyəlik səs yazısı ilə şəxsin səsində istənilən mətni oxuyaraq həmin şəxsin süni səsini yaradır;
- Runway (Gen-3) - videodakı hər hansı obyekti və ya insanı süni şəkildə dəyişərək, arxa fonu tamamilə manipulyasiya edə bilir;
- Reface/FaceApp - əyləncə məqsədli üz dəyişdirmə. Məsələn, şəxsin üzünü məşhur bir film qəhrəmanının bədəninə "köçürür".

İkinci Qarabağ müharibəsindəki parlaq qələbəmizdən sonra yeni texnologiyaların, eləcə də, süni intellektin tədqiqi və tətbiqi ölkəmizdə geniş vüsət almışdır. Azərbaycan Respublikası Prezidentinin Sərəncamı ilə təsdiq olunmuş "Azərbaycan Respublikasının 2025–2028-ci illər üçün süni intellekt Strategiyası"nın Tədbirlər Planı bu sahədə istiqamətverici və aparıcı sənədlərdən birinə çevrilmişdir.

Yuxarıda qeyd etdiyimiz alət və vasitələrdən istifadə ilə milli proqram məhsullarının yaradılması istiqamətində tədqiqatlar və praktiki xarakterli işlər davam etdirilir. Azərbaycan Respublikasının Xüsusi Rabitə və İnformasiya Təhlükəsizliyi Dövlət Xidmətinin əməkdaşları tərəfindən hazırlanmış saxta bir video nümunəsinin yaradılması alqoritmi diqqətinizə çatdırıram [9]:

1) Hədəf şəklin seçilməsi - ilk olaraq "source_image" (mənbə şəkli) hissəsinə Deepfake tətbiq olunacaq şəxsin (məsələn, Barak Obamanın) fotoşəkili yüklənir;

2) Hərəkətverici ("Driving") videonun seçilməsi - "driving_video" hissəsinə hədəf şəkli hərəkətə gətirəcək, mimikaların və baş hərəkətlərinin götürüləcəyi video (məsələn: Jackie Chan) əlavə edilir;

3) Dataset-in təyin edilməsi - proqram daxilində uyğun dataset (məsələn, VOX) seçilir. Bu, süni intellektin üz hərəkətlərini daha dəqiq analiz etməsinə kömək edir;

4) Prosesin işə salınması - "Run" düyməsinə kliklədilməklə sintez prosesi başladılır;

5) "Deep Learning" analizi - proqram dərin öyrənmə texnologiyasından istifadə edərək şəklin və videonun rəqəmsal parametrlərini (üz cizgiləri, nöqtələr və s.) analiz edir;

6) Parametrlərin hazırlanması - hərəkətverici videodakı hərəkətlər rəqəmsal parametrlərə çevrilir;

7) Sinxronlaşdırma - hazırlanmış parametrlər hədəf şəkil ilə üst-üstə qoyularaq sinxronlaşdırılır;

8) Nəticənin ekrana çıxarılması - sintez olunmuş yeni video "Output" hissəsində nümayiş etdirilir;

9) Videonun yaddaşa verilməsi - son addımda hazırlanmış Deepfake videosu "Download" düyməsi vasitəsilə kompüterə yüklənir.

IV. SÜNİ İNTELLEKT AGENTLƏRİ

İlk növbədə süni intellekt agenti nədir sualına cavab vermək lazımdır. Bilirik ki, real həyatın bütün sahələrində hər hansı bir məsələnin həlli zamanı qərar verən insanlar öz məqsədlərinə uyğun olaraq, ilk növbədə aidiyyəti verilənləri toplayır, analiz edir və bu verilənlərə əsasən qərar qəbul edir.

"Süni intellekt kompüterlərin insan beynində gedən düşünmə və öyrənmə proseslərini təqlid etməsidir" desək yanılmarıq. Başqa sözlə, süni intellekt kompüter sistemlərinin və proqramların insan zəkasına xas olan öyrənmə, məntiqi qərarvermə, problemlərin həlli, nitqi və təsvirləri tanıma kimi bacarıqları icra edə bilməsini təmin edən texnologiyalar və üsullar toplusudur. Qeyd etdiyimiz bacarıqların süni intellekt vasitəsilə icrasını təmin edən əsas alət Sİ agentləridir. Sİ agentləri ətraf mühiti qavrayan, məlumat toplayan və daimi nəzarət olmadan insan tərəfindən müəyyən edilmiş məqsədlərə çatmaq üçün müstəqil qərarlar qəbul edən avtonom proqram sistemləridir.

Məqsədlər insanlar tərəfindən müəyyən edilir və Sİ agenti onlara çatmaq üçün müstəqil olaraq optimal hərəkətləri seçir. Məsələn, müştəri sorğularını həll etmək istəyən əlaqə mərkəzindəki Sİ agentinin fəaliyyəti belədir: agent avtomatik olaraq müştəriyə suallar verməli, daxili sənədlərdə məlumat axtarmalı və həll yolu ilə cavab verməli olacaqdır. Müştərinin cavablarına əsasən, problemi müstəqil şəkildə həll edə biləcəyini və ya sorğunu bir insana həvalə etməli olub-olmadığını müəyyən edə bilir.

Süni intellekt agentlərini xarakterizə edən əsas prinsiplər bunlardır [10]: avtonomluq; məqsədyönlü davranış; qavrayış; rasionallıq; uyğunlaşma; davamlı öyrənmə; birgə iş.

Sİ agentlərinin işləmə prinsipi insan düşüncə prosesinə bənzər alqoritmik şəkildə belə baş verir:

1. *Məlumat toplama (Perception);*
2. *Analiz və qərarvermə (Processing);*
3. *Fəaliyyət (Action);*
4. *Öyrənmə (Learning).*

Dünyada geniş şəkildə istifadə olunan Sİ agentlərinə aşağıdakı nümunələri göstərmək olar:

- Devin (Cognition AI);
- AutoGPT/BabyAGI;
- Agentforce (Salesforce);
- Microsoft Copilot (Pro versiyası);
- ChatGPT (OpenAI);
- Google Gemini (Ekosistem ustası);
- DeepSeek (Məntiq və qiymət çempionu).

V. SÜNİ İNTELLEKT ƏSASLI DEZİNFORMASİYA İLƏ MÜBARİZƏ

Deepfake texnologiyası sürətlə inkişaf edir. Cinayətkarlar da daxil olmaqla, hər kəsin tapa biləcəyi əlamətlər hələ də mövcuddur. Bunlara aşağıdakıları aid etmək olar [8]:

- Qeyri-təbii göz hərəkəti və ifadələri - gözlərin həddindən artıq və ya kifayət qədər qırpmadığının, yaxud üz ifadələrinin səs tonuna uyğun olmadığının müşahidə edilməsi;
- Nitqdə yöndəmsiz temp və ya emosiya - səsdə qeyri-təbii fasilələrə, qəribə intonasiyalara və ya robotik emosional diapazon çatışmazlığının müşahidə edilməsi;
- Bulanıq və ya uyğun olmayan kənarlar - üzün saçla, boyunla və ya ağız ətrafında birləşdiyi yerdə rəqəmsal artefaktların, bulanıqlığın və ya qəribə keçidlərin olması. Üzün saçla və ya boyunla birləşdiyi yerlərdə və ağız ətrafında rəqəmsal artefaktların, bulanıqlığın və ya qəribə keçidlərin olması;
- Zəif dodaq sinxronlaşması - səsin ağız hərəkətləri ilə mükəmməl sinxronlaşdırılmaması;
- Uyğunsuz vizual və audio keyfiyyəti - video və audio keyfiyyətinin fərqli olması və qeyri-adi fon səslərinin mövcudluğu;
- Qeyri-adi vadi effekti - deepfake bəzən demək olar ki, mükəmməl görünür, lakin yenə də müəyyən etmək çətin olan halda "səhv" və ya müəyyən narahatlıq hissinin olması. Bu instinktə etibar edilməlidir;
- Doğrulama/verifikasiya protokolu - hər hansı qeyri-adi və ya yüksək riskli sorğu üçün həmişə ayrı, etibarlı rabitə kanalı vasitəsilə yoxlamanın vacibliyi.

Unutmaq olmaz ki, cinayətkarlar saxta sorğularını daha inandırıcı etmək üçün tez-tez açıq mənbələrdən toplanmış məlumatlardan istifadə edirlər və bu da qrupdan kənar yoxlamanı daha da vacib edir.

"Deepfake" hücumuna qarşı mübarizədə atılası vacib addımlar bunlardır:

- Görüntülərə diqqətlə baxmaq lazımdır. Video və ya şəkildə görünən şəxsin üz təsvirinə və ümumi keyfiyyət kimi detallarına diqqət yetirilməlidir;
- Cavab strategiyası hazırlanmalıdır. İstifadəçi deepfake-ə düzgün şəkildə reaksiya verməyə hazır olduğundan əmin olmalıdır;
- Yeni kibertəhlükəsizlik standartları yaradılmalıdır. Misal olaraq "əməkdaşların kritik əməliyyatları həyata

keçirməzdən əvvəl onun reallığını təsdiqləmək üçün addımlar atmasını” göstərmək olar;

- Əməkdaşların və ümumiyyətlə cəmiyyətin bütün üzvlərinin maarifləndirilməsi müntəzəm olaraq həyata keçirilməlidir.
- Mütəmadi olaraq kibertəhlükəsizlik və süni intellekt maarifləndirmə təlimləri həyata keçirilməlidir;
- Unutmaq olmaz ki, “deepfake” və digər kiberhücumlar ilə mübarizə məlumatlı olmaqdan başlayır.

Cəmiyyət demək olar ki, hər gün ortaya çıxan yeni bir problemlə üz-üzə qalır və təbii olaraq həmin problemlərlə mübarizə immuniteti də yaranmaqdadır. Bu cur zərərli vasitələrə qarşı ən yaxşı mübarizə isə maarifləmə, diqqət və texnoloji bilgilərdir.

Cəmiyyəti yalnız rəsmi mənbələrin verdiyi informasiyaya inanmağa, jurnalistləri, ictimai fəalları bu kimi hallara qarşı hər zaman prinsipiallıq nümayiş etdirməyə, saxta və yalan məlumat əsaslı kampaniyaların, “deepfake” texnologiyaları ilə məzmunyaratma tendensiyasının vüsət aldığı bir şəraitdə sayıq olmağa səsləyirik.

Məşğuliyyətindən asılı olmayaraq hər bir ölkə vətəndaşı aşağıdakı tövsiyələrə əməl etməlidir:

- Hər görünən video və eşidilən səsə dərhal inanmaq olmaz. Xüsusilə, “təcili” xarakterli çağırışlar, pul və ya şəxsi məlumat tələbləri şübhə ilə qarşılanmalıdır;
- Məlumatı alternativ mənbələrdən yoxlamaq vacibdir. Əhəmiyyətli bir xəbər və ya açıqlama yalnız bir mənbədə varsa, onun doğruluğu sual altındadır. Etibarlı media və rəsmi kanallar əsas olmalıdır;
- Rəsmi qaynaqlarla işləmək vacib amildir;
- Vətəndaşların rəqəmsal savadlılığı davamlı artırılmalıdır. Çünki, belə savadı olan insan manipulyasiyanı daha tez anlayır, emosional qərarlar vermir və informasiya təhlükəsizliyinə daha məsuliyyətlə yanaşır;
- Rəqəmsal savadlılığa ailədən, uşaqlardan, məktəblərdən başlamaq lazımdır. Uşaqların deepfake və ümumilikdə rəqəmsal saxtakarlıqlar barədə məlumatlandırılması müasir təhsil sisteminin vacib tərkib hissəsinə çevrilməlidir;
- Bütün tədris müəssisələrində şagirdlərin və tələbələrin yaş səviyyəsinə uyğun rəqəmsal savadlılıq dərsləri, real və saxta məzmunu ayırd etməyə yönəlmiş praktik nümunələrin nümayişi, həmçinin onlayn təhlükəsizlik və şəxsi məlumatların qorunması mövzusunda maarifləndirici dəyirmi masalar təşkil olunmalı və keçirilməlidir;
- Müəllimlər şagirdlərə hər gördükləri video və ya eşitdikləri məlumata tənqidi yanaşmağı, şübhəli hallarda valideyn və ya müəllimə müraciət etməyin vacibliyini aşılamalıdırlar.

Bir məsələni xüsusi olaraq vurğulamaq istəyirəm. İnformasiya təhlükəsizliyi, kibertəhlükəsizlik və süni intellekt sahəsində yüksək səviyyədə nəzəri biliklərə və praktiki bacarıqlara malik mütəxəssislər yetişdirilməlidir. Bu məqsədə çatmaq üçün ölkədə kibertəhlükəsizlik, süni intellekt, rəqəmsal kriminalistika, kriptoqrafiya və s. praktiki laboratoriyaların yaradılması vacibdir.

NƏTİCƏ

Real görüntü, səs və video ilə onların deepfake ilə yaradılmış rəqəmsal forması arasında hər hansı bir fərq bulanıqdır. Dərin bir hücuma cavab gözləmək, əksər təşkilatların ödəyə bilmədiyi bir riskdir. Lakin belə risklərin idarə edilməsi və minimum endirilməsi üçün qabaqlayıcı tədbirlər, davamlı təlim və təhlükəsizlik maarifləndirmə mədəniyyəti ən təsirli strategiyalardır. Bütün dünyada yeni təhdid növü kimi qiymətləndirilən süni intellekt əsaslı dezinformasiyalara qarşı ölkəmizdə müvafiq qurumlar tərəfindən mübarizə və geniş maarifləndirmə işləri aparılmaqdadır. Nəticə etibari deepfake texnologiyasının inkişafı qarşısında tam müdafiə mümkün olmasa da, məlumatlılıq, tənqidi düşüncə və düzgün davranış vərdişləri bu riskləri minimuma endirə bilər. Cəmiyyət olaraq əsas vəzifə yalnız texnologiyanı izləmək deyil, onu düzgün istifadə etməyi də öyrənməkdir.

Qeyd edək ki, yaxın gələcəkdə mənbə materiallarının Açıq mənbə kəşfiyyatı (OSINT) vasitəsilə asanlıqla əldə edilməsi ilə Deepfake texnologiyasının gücləndirilməsi və nəticədə təhdidlər sahəsində əhəmiyyətli bir təkamülün də baş verməsi qaçılmazdır. Yaradıcı və faydalı tətbiqlər təklif edilsə də, fırıldaqçılıq, dezinformasiya və nüfuza zərər vurmaq üçün zərərli istifadə potensialı danılmaz olaraq qalacaqdır. Süni intellekt əsaslı dezinformasiyaya qarşı mübarizə sahəsində yuxarıda qeyd etdiyimiz tədbirlərlə yanaşı mütləq mənada milli proqram təminatlarımızın (süni intellect agentlərinin, maşın öyrənməsi, kriptoqrafiya və digər vacib istiqamətlər üzrə “millilləşdirilmiş” alqoritmlərin, platformaların və s. alət və vasitələrin) yaradılması istiqmətində fəaliyyətimizi daim gücləndirməliyik. İnanırıq ki, bu istiqamət daim dövlətimizin diqqət mərkəzindədir və bu sahəyə lazımi dəstəyi verəcəkdir.

ƏDƏBİYYAT

[1] “Saxta Xəbər” nədir? https://guides.lib.umich.edu/fakenews

[2] https://www.dw.com/ru/пять-самых-популярных-фейков-о-европе-в-русскоязычных-соцсетях/a-47501922.

[3] https://www.merriam-webster.com/dictionary/disinformation.

[4] Фейки недели. https://investigatebel.org/ru/fakenews/07-09-2023-top-5-fake-news.

[5] Deepfake. https://www.unit21.ai/fraud-aml-dictionary/deepfake.

[6] Анкара направит 10 миллионов долларов «для поддержки уязвимых слоев армянского общества». https://geworld.ge/ru/анкара-направит-10-миллионов-долларов/.

[7] https://minval.az/news/124513351.

[8] 5 Deepfake Examples of AI Videos & Images That Blur Reality. https://www.adaptivesecurity.com/blog/deepfake-examples-ai-videos-images.

[9] https://scis.gov.az/

[10] Что такое агенты искусственного интеллекта? https://aws.amazon.com/ru/what-is/ai-agents/

Creation, Application and Combating Artificial Intelligence-Based Disinformation

Ziyafat Amirov

The Academy of the State Security Service of the Republic of Azerbaijan named after Heydar Aliyev, Baku, Azerbaijan

Abstract— In the context of the global information wars of the modern era, the spread of fake information and news has become one of the main tools for manipulating public consciousness. Fake information poses a threat to national security and cybersovereignty, undermines trust in state institutions, and causes various problems in the social environment. The issue of creating, using and effectively combating fake information using artificial intelligence is becoming relevant. The article analyzes the concepts of fake news, deep fake (deepfake), artificial intelligence agent, as well as the mechanisms for creating artificial intelligence-based disinformation, its application in the cybersecurity environment and the risks it poses. The technical and organizational defense approaches used to detect disinformation, the role of media literacy, and legal regulations are commented on.

Keywords— national security; digital sovereignty; cybersovereignty; disinformation; fake news; deep fake; artificial intelligence agent.