A MESTERSÉGES INTELLIGENCIA ALAPÚ KIBERTÁMADÁSOK LEHETŐSÉGEI

THE POSSIBILITIES OF THE ARTIFICIAL INTELLIGENCE-BASED CYBER-ATTACKS

Fehér András Tibor[1], Négyesi Imre[2]

[1]egyetemi tanársegéd, [2]egyetemi docens
[1-2] Nemzeti Közszolgálati Egyetem Hadtudományi és Honvédtisztképző Kar
E-mail: feher.andras@uni-nke.hu [1], negyesi.imre@uni-nke.hu [2]

Összefoglalás

A számítógépek, az okoseszközök és az internet fejlődése létrehozta az úgynevezett kiberteret, amit számos előnye mellett a bűnözői, katonai vagy politikai erők is igénybe vehetnek az emberek vagy az emberiség ellen. A modern technológiák csúcsa, a mesterséges intelligencia is számos veszélyt rejt, épp a benne rejlő lehetőségek miatt. Fő célunk e két újszerű veszélyforrás keresztmetszetének bemutatása. Ez által egyben választ kaphatunk olyan kérdésekre is, hogy hogyan kell hozzáállnunk ehhez a veszélyforráshoz nekünk, mint polgárnak, mint vezetőnek, vagy mint informatikai üzemeltetőnek.

Abstract

The development of computers, smart devices, and the Internet has created what is known as cyberspace, which, in addition to its many benefits, can also be used by criminal, military, or political forces against people or humanity. The pinnacle of modern technology, the artificial intelligence, also holds many dangers precisely because of its potential. Our main goal is to present the cross-section of these two novel hazards. At the same time, we can get answers to questions about how we have to approach this source of danger, as a citizen, as a leader, or as an IT operator.

Kulcsszavak: *mesterséges intelligencia, kibertér, kibertámadás*
JEL besorolás: *D82*
LCC kód: *T173.2-174.5*

Bevezetés

Hogy jelen cikk olvasásának motivációját megadjuk, kezdjük kutatás jelentőségének hangsúlyozásával. A mesterséges intelligenciának (továbbiakban MI) számtalan felhasználási lehetősége létezik. Hazánk MI Stratégiája (ITM, 2020), az Integrált, digitális agráriumtól, a logisztikán át, az adatvezérelt egészségügyig közérthetően vázolja legfontosabb területeket. Azonban minden technológia veszélyessé válhat, mégpedig kétféleképpen: vagy a békés célú felhasználás romlik el (pl. repülőgép baleset, atomerőmű-baleset) vagy kifejezetten támadó célra fejlesztenek tovább egy találmányt (pl. vadászrepülőgép vagy atombomba). Az alábbiakban a második esetre koncentrálunk. Ezen belül azonban elmondható, hogy az MI a hadiipar legtöbb területén még csupán prototípusok, kutatások, tesztelések fázisában jár (Négyesi, 2017). Ezzel szemben az MI kiber-támadó képességének jelentőségére egyre több külföldi szakmai bemutató és publikáció hívja fel arra a figyelmet, már 2016 óta.

 10.18531/Studia.Mundi.2021.08.05.25-34

Általánosságban elmondható, hogy a szakemberek egyetértenek abban, hogy a fenyegetések új generációjával állunk szemben, amely át fogja írni az eddig megszokott védelmi elveket is. A legkifejezőbb megfogalmazás (Rodriguez, 2019) szerint az MI miatt a kiberbiztonság helyzete most hasonló lett a 90-es évek közepének kihívásaihoz. A kérdés most az, hogy a számítógépek tanuló és döntéshozó képességét kik tudják hatékonyabban kihasználni: különféle állami, alvilági, terrorista stb. kibertámadó rendszerek – vagy a kibervédelem. Tehát nem csiszolgatni kell a védelmet, hanem teljesen új alapokra kell helyezni. Ehhez a paradigmaváltáshoz szeretnénk hozzájárulni, és felhívni a figyelmet erre a területre, mivel hazánkban erről alig jelent meg tanulmány. Reméljük, hogy az itt összefoglalt ismeretek alapul szolgálhatnak egy ilyen irányú kockázat-elemzés kidolgozásához is. Eközben nem szándékunk senkit megijeszteni, ezért nem maradunk végül adósak annak megvizsgálásával sem, hogy vajon tényleg sürgető-e ez a probléma? Milyen hozzáállást indokolt ezen a téren magunkban kialakítanunk olyan emberekként, aki felhasználóként kapcsolódik az otthoni vagy céges számítógéprendszerekhez? Az MI veszélyeivel kapcsolatos hozzáállásnak két véglete van:

(1.) Akik az MI képességeit látva könnyebben „farkast kiáltanak", s így rémhírek keletkeznek szakmai körökben is. A rémhírkeltésre legjobb példa, hogy 2017-ben a Black Hat konferencián részt vevő szakemberek között egy felmérést végeztek (ISSSource, 2017), melynek eredménye szerint 2018-ra a kibertámadások 62%-a mögött a mesterséges intelligencia fog állni. Ez a nézet akkor elterjedt, mivel hírt igen sok hírportál átvette – ám máig nem teljesült be, még 2021 elején sem található erre utaló adat (Sobers, 2020).

(2.) A másik véglet, hogy akik pedig megszokták „okos-dolgaik" ártalmatlanságát, azok hajlamosabbak lehetnek elbagatellizálni a velük kapcsolatos (esetleg áttételes) veszélyeket. A kételkedők egyik fő érve, hogy az MI költséges dolog, ezért lassan fog terjedni. Ám a tények sajnos ezt megcáfolják, hiszen jócskán van pénzük a támadóknak ilyen drága technológiára is. Egyrészt nem zárhatóak ki jelentős állami (hadi) támogatások, másrészt óriási bevétele van a különféle csoportoknak kiberbűnözésből. Már 2017-ben 1,5 billió (ezermilliárd, 10^{12}) dollárra volt becsülhető (McGuire, 2018) a hozzájuk befolyt összeg, ami 2021-re elérheti a 6 billiót is (Morgan, 2019). Ám ezek az érvek-ellenérvek valójában irrelevánsak, nem az érvelés miatt félnek kevesen ettől a veszélytől, hanem mivel a többség nem tud róla, vagy nem vesz tudomást róla.

Ha összegezzük a félelem és a kételkedés fenti vizsgálatait, az rajzolódik ki, hogy egyik oldalon sem igazi érvek állnak szemben: akik informatikával foglalkoznak, jobban hajlanak arra, hogy féljenek a kibertámadó MI-től, de nekik csak rémhíreik vannak – a többiek könnyen negligálják a problémát, ebben főleg a nemtudás játszik szerepet. Ezekre a hozzáállásokra a végkövetkeztetéseknél térünk majd vissza.

Anyag és Módszertan

Metódusunk alapvetően elemző-rendszerező módszerekre épül, konkrétabban az elmúlt néhány év civil MI technológiáinak helyzetelemzésére. (Kutatásunk kezdetén a katonai forrásokat is át kívántuk tekinteni, de a nyilvános katonai anyagokban az MI alapú kibertámadási technikákról legfeljebb általánosságban esik szó, ezért ezt a kutatási irányt elvetettük.) Az összegyűjtött információkat nyolc pontba rendeztük. Mindegyiknél megvizsgáljuk, hogy az MI megjelenése milyen új lehetőségeket vetett fel a kibertámadások szempontjából, valamint, hogy azt a lehetőséget kihasználva mennyire alkalmas az MI egy nagyerejű kibertámadásra. Mindezt összegezve válik láthatóvá, hogy közvélekedés mennyire reagálja túl vagy alul a helyzetet, mennyiben tér el egy optimális, kockázat-elemzésre alapuló állásponttól. Az összegzésből

 10.18531/Studia.Mundi.2021.08.05.25-34

pedig diszkurzív következtetéssel juthatunk reális javaslatokhoz arra nézve, mit tegyen a problémával kapcsolatban egy IT szakember, vagy egy vezető.

Eredmények

Itt érkeztünk el oda, hogy a kibertérben támadó mesterséges intelligencia főbb lehetőségeit megvizsgáljuk. A fellelhető szakirodalom alapján az alábbi nyolc pontba volt célszerű rendezni a fellelt anyagot. Ezek összessége megfelelő áttekintést ad azokról a legfontosabb veszélyforrásokról, ahol az MI lehetőségei kihasználhatók.

1. A megnövekedett digitális támadási felület kihasználása

Ahogyan nem mindegy, hogy egy sziklaszorost kell védeni adott létszámú emberrel, vagy egy több száz kilométeres partszakaszt, úgy a technikai bonyolultság is egy ilyen nagyon hosszú partszakasszá válhat: körülbelül ezt foglalhatjuk össze abban a kifejezésben, hogy „egyre nagyobb a digitális támadási felület". Más szóval egyre több a potenciális sérülékenység minden rendszerünkben, eszközünkben, programunkban. A technológiák az emberek által teljesen áttekinthetetlenné váltak. Egy adott rendszer téves beállításait vagy egyéb hibáit eddig is keresték a támadók vírusai, ám azok csupán konkrét hibák után voltak képesek kutatni. Kétféle megközelítési mód volt használatos: az egyikben ismert hibatípusokhoz hasonlókat kutatnak fel, a másikban a felhasználói hibajelenségek elemzésével találják meg a sérülékenységet. Mivel az MI használata nélkül a támadók is csupán a tesztelő programok és kódelemzés dimenzióiban mozoghattak, ezért egyfajta szerencse is kellett, hogy olyan rést találjanak, amit a védelmi oldal nem derített fel. Az MI azonban szintet lép a kódelemzésben, mivel a lehetséges problémák mintázatait keresi – így akár milliószor több hibalehetőség lelhető meg. Ennek érzékeltetése egy emberi példán célszerű: egy rendőrségi körözés alkalmával egy konkrét arcot keresnek, ez felel meg a hagyományos algoritmusok keresésének – ám amikor ez nem áll rendelkezésre, a rendőrök például gyanús viselkedések beazonosításával próbálják a lehetséges elkövetők körét leszűkíteni. Vannak bizonyos jelei annak amit „gyanúsnak" hívunk, több apró, külön-külön érdektelen gesztus, viselkedés, reakció, és a jelek bizonyos összefüggése (mintázata) helyez egy megfigyelt személyt a „gyanús" halmazba. Ez a módszer állítható párhuzamba az MI-t használó, minta alapú kereséssel. Persze a gyanúsan viselkedők egy része nem a köthető bűntényhez, épp így az MI által megtalált résnek tűnő helyek sem feltétlenül jelentenek behatolási pontot a támadott rendszerbe. Tehát a megtalált potenciális réseken a támadó MI próbálkozik egy sor tipikus módszerrel, végül jelenti egy emberi szakértőnek mire jutott, sikerült-e valamely próba, és miért találja mégis gyanúsnak a kapott lehetőséget. A szakértő ezután megmondja az MI-nek, hogy az valóban rés volt-e avagy sem, így miközben a támadott rendszer sérülékeny pontjait is beazonosítja, egyúttal az MI-t is tanítja. A példánknál maradva, ahogyan egy idősebb, tapasztaltabb rendőr sokkal kisebb dolgokat észrevesz és azonosít be gyanúsként – ugyanígy egy MI is fejlődhet, és a szakember által tanítva egyre több potenciális rést vesz észre.

Ezzel a módszerrel tehát a hackerek támadás-előkészítési képessége erősödött meg nagyon, vagyis a konkrét támadást az MI által felderített gyenge pontokra koncentrálva hajtják végre. Egy ilyen offenzíva tehát úgy képzelhető el, hogy egy MI modul végzi el a felderítési munkát, más MI modulok pedig a támadási feladat végrehajtásánál használják fel a kapott információt (Chartis Research, 2019). Ha ehhez hozzávesszük, hogy az MI más területeken (álcázás, célratartás, stb.) is igen sikeres (Ph. Stoecklin - Jiyong - Kirat, 2018), tetszőleges számú modullal kiegészítve a módszert, a támadás hatékonysága még tovább növelhető. A felsorolt veszélyeket összegezve elmondható, hogy a támadók előnye fokozatosan növekszik, sőt az elavult rendszerekkel szemben már most számottevő. Ezt kompenzálja, hogy a szoftver és a

 10.18531/Studia.Mundi.2021.08.05.25-34

hardver gyártók védelmi szakemberei is már jó ideje az MI-t hívják segítségül a támadási felület átlátásához, ám a kisebb vállalatok vagy magánemberek egyelőre csak a frissítések vagy vírus-adatbázisok formájában jutnak hozzá az MI által felderített és javított rések foltozásához. Viszont éveken belül szükségessé válhat közepes méretű hálózatoktól kezdve saját tanítású MI modulok bevezetése az informatikai védelmi rendszerekben.

A digitális támadási felülethez szorosan kapcsolódik, hogy a megtalált sérülékenységek kezelését már régóta különféle adatbázisok segítségével oldják meg (például CVE, NVD, US-CERT, Snyk Vulnerability adatbázisok). Ezekben az adatbázisokban tárolják a problémákat és azok javításait, a rengeteg program és eszköz biztonságosan tartásához. Azért van rájuk szükség, mert rések nem csupán a rendszer létrehozása során keletkeznek, hanem valamely részösszetevő telepítéséből, törléséből vagy frissítéséből, egyszerűen a rendszer elhasználódásából is adódhatnak. Ez korábban azért nem jelentett veszélyt, mivel egy folyamatosan frissített rendszer nem tartalmazott az adatbázisban tárolt rést. (A rosszul karbantartott rendszerek támadására eddig is lehetett ezeket az adatbázisokat használni, de a frissítést elmulasztó felhasználók felé a gyártók soha nem is vállaltak garanciát.) Ám az, hogy ezeket az adatokat offenzív MI rendszerek betanítására is fel lehet használni, az új jelenség (Fu, 2018). Ugyanis, ha egy fejlettebb MI képes mintákat felfedezni az adatbázisban, akkor képes lesz felismerni az adatbázisban szereplőhöz hasonló, de eddig ismeretlen réseket is egy rendszer a vizsgálatakor (felhasználva a talált mintát, ahogyan azt az 1-es pontban bemutattuk).

Összegzésként: ez a lehetőség sem ad okot arra, hogy a veszélyt eltúlozzuk, hiszen a védelmi szakembereknek talán többet segít az MI a rések foltozásakor. Felvetődik viszont, hogy újra kell gondolni minden publikus adatbázis használatát – köztük a sérülékenységi adatbázisokét is – hogy hasonló, illegitim, hacker-célú MI-k betanítását megakadályozzuk. Ez az újragondolás ugyan már megkezdődött, elsősorban a nagy számú automatikus lekérdezés elleni védelmi módszereket vezettek be, de ez hosszú távon nem lesz elegendő, szükség lesz az adatbázisok használóinak engedélyezési és hitelesítési eljárásrendjét nemzetközileg rögzíteni.

3. Az IoT és az Ipar 4.0 veszélyei

Olyan korszak kezdetén állunk, ahol IoT (Internet of Things – a „dolgok” internete), vagyis a hálózatba kötött eszközök és szenzorok lesznek hivatottak még hatékonyabbá tenni termelési folyamatokat, és még kényelmesebbé tenni életünket. Ez a sok digitális „dolog” rengeteg adatot szolgáltat, így keletkeznek a BigData-nak nevezett, a korábbi adatbázisokat nagyságrendileg meghaladó adatbázisok, amelyek lehetővé teszik a nagyobb hatékonyságot. Ez megköveteli az MI használatát, mivel ilyen nagy mennyiségű adat emberek által nem dolgozható fel, az MI tanításához viszont épp ez az óriási adatmennyiség kívánatos (sőt épp ennek hiánya volt az egyik oka, hogy nem terjedt el korábban). Az IoT, a BigData, az MI és az új, nagy sávszélességű gyors hálózati technológiák (pl. OC768 vezetékes, wifi 6 vezeték nélküli vagy 5G mobiltelefon szabványok) együttesen teszik kényelmesebbé az egész világot: a személyes életünket, a szolgáltatásokat, a társadalmat, a politikát, tehát mindent. Csakhogy ezzel együtt támadhatóbbá is tesznek mindent.

Az IoT által közeledik egymáshoz a fizikai és digitális világ. Így a kibertámadás már nem csupán számítógépet bénít, adatot manipulál, hanem a rosszindulatú programok megfertőzhetik okos-eszközökkel teli fizikai valóságot is: hibás időt mutat az okosóra, leáll az okosautó, rosszul rendel az okoshűtő. Igaz ez az internethez csatlakozó daruktól kezdve, az utcai lámpák érzékelőin át, a gyári szállítószalagokig és a szélturbinákig – akár nem is fő célpontként, csupán járulékos, de súlyos kárként. Az IoT-eszközök is hozzájárulnak a fentebb említett a támadási felületet kiszélesedéséhez (ld. 1.). Nagyban épül az IoT technológiára is az úgynevezett Ipar

4.0. Ez a termelési folyamatok olyan szervezését írja le, melynek keretében az eszközök önállóan kommunikálnak egymással az értéklánc mentén: a jövő egy olyan „okos" gyárát hozva létre ezzel, amelyben a számítógép-vezérelt rendszerek nyomon követik a fizikai folyamatokat, létrehozzák a fizikai valóság virtuális mását és decentralizált döntéseket hoznak önszervező mechanizmusok alapján (ITRE, 2016). A fő problémát itt az jelenti, hogy az Ipar 4.0 nem csak vadonatúj egységekből épül fel, hanem a régi rendszereket látják el IoT kiterjesztéssel, vagyis olyan elavult rendszerekre támaszkodik, melyeket nem erre a korszakra terveztek. Ennek következtében olyan számítógépes sebezhetőségekkel vannak tele, amelyeket majd csak egy támadás után fedeznek fel, amikor már késő. Az intelligens városokoknak és az ipari környezeteknek tehát újra kell gondolniuk egész kibervédelmüket. Ez olyan terület, aminek gyengeségeire alapozhatja támadásait akár egy ellenséges ország is (Tech Wire Asia, 2020). Hozzá kell tennünk, hogy az otthoni IoT-eszközök, is veszélyesek, mivel hírhedten rossz biztonságuk miatt jól használhatóak például elosztott szolgáltatásmegtagadással járó támadások (DDoS, Distributed Denial of Service) során (Dickson, 2020).

Mindent egybevetve talán ez a legérzékenyebb terület kibervédelmi szempontból az itt tárgyalt nyolc közül, viszont nem kezelhető azokkal együtt, mivel nem a támadó MI-re épül. Nem az IoT „okossága", hanem inkább „butasága" veszélyes (elavult, olcsó vagy rosszul beállított elemei). Ezek a rések hagyományos módszerekkel talán sokkal egyszerűbben felderíthetőek és támadhatóak.

4. Az mesterséges intelligencia és BigData léte kihasználható

Nem csupán úgy lehet veszélyes az MI, hogy felhasználják a támadás végrehajtásakor. A különböző rendszereket épp az veszélyezteti többféle módon is, hogy MI-t használnak. Az első mód, ha a kibertámadó megszerzi vagy visszafejteni azt a BigData adathalmazt, amivel a saját (tetszőleges célú) MI rendszerünket betanítottuk. Ezeket az adatokat sok féle módon fel lehet használni visszaélésekre: képet adnak a támadónak rólunk, felhasználhatja saját MI rendszerének betanításához, de klasszikus kibermódszerekhez is alkalmazhatóak, például jelszótöréshez személyes adatokból tippeket kaphat belőle (Durbin, 2020). Kibertéren kívül is használhatóak, például támadó az adatbázisban szereplő érzékeny adatokat egy-egy prominens személy zsarolására. A második mód, ha az egész betanított MI-t lopják el és fejlesztik tovább, fordítják ellenünk. Harmadik mód, ha nem ellopják az MI-t, hanem a tanítási folyamatot törik fel és mérgezik meg, finoman kiképezve az algoritmust a támadó céljainak megfelelő döntésekre (Da Silva, 2019). Egy ilyen jól megtervezett tanulás-fertőzéssel elérhető, hogy a rendszer néhány irányzottan rossz választ hozzon, a támadó szándéka szerint manipulálva az MI elemzésére épülő döntéseket. Az adatbázis, vagy az MI kód megszerzésének veszélyessége nem kiemelkedő, mivel az előző ponthoz hasonlóan inkább a klasszikus kibertámadási módszerekkel valószínű – ennek megfelelően ez sem sorolható szorosan a témánkhoz.

5. A támadások kifinomultabbak és személyre szabottak lesznek

A beszéd és a szokások mesterséges tanulhatósága, valamint a képfelismerés és képalkotás fejlettsége új távlatokat nyitott meg a modern szélhámosságban, amit úgy hívunk *social engineering* (más kifejezéssel pszichológiai befolyásolás). Szempontunkból kétfelé oszthatjuk ezt a digitális csalásformát. Az alábbi felosztás a manapság használatos MI alapú támadásokon alapul, eltérően a szakirodalom egyéb felosztásaitól (Muha - Krasznay, 2018). Az MI alapú social engineering egyik része a valóság digitális utánzásán alapul, nevezzük „utánzó szélhámosság"-nak. Ide tartoznak a deepfake technikák (MI technológiák segítségével meghamisított videók), vagy a digitális megszemélyesítés. Az MI alapú social engineering másik részénél az emberek személyes, egyedi szokásrendszerének egyre tökéletesebb

 10.18531/Studia.Mundi.2021.08.05.25-34

feltérképezése a kulcs, nevezzük „kutató szélhámosság”-nak. Ez a szegmens hasonló ahhoz, amit a máshol például „a social engineering információgyűjtő fázisa”-ként említenek (Kovács, 2018).

A bizalom megszerzése a szélhámosok ősi alaptrükkje, ennek MI alapú megvalósítását a kutató szélhámosság esetében egy elképzelt (és elképzelhető) spam alapú példával szemléltetjük. Azzal mindenki találkozott már, hogy valamilyen ürüggyel egy adathalász weboldalra, vagy vírusos melléklet megnyitására, netán adatainak megadására akarják rávenni. Sokan találkozhattak már chat-robotokkal is, melyet például ügyfélszolgálatok használnak a tipikus problémák automatikus megválaszolására (ezek is MI-t használnak). A kettő egyesítésével létrehozható olyan chat-bothoz hasonló, tanítható program, amely emailekre válaszol automatikusan és „intelligensen”. Ennek előnye, hogy nem szükséges az első alkalommal letámadni a célszemélyt – hiszen egyre kevesebben dőlnek be ezeknek a próbálkozásoknak. A spam-robot viszont képes néhány levélváltással kialakítani egy kis bizalmat. Ezután az áldozat gyanútlanul besétál a „kedves ügyfélszolgálatos” (valójában az MI) által küldött csapdába. Így adatokat is ki lehet csalni az áldozatból, aki azt hiszi valós cégnek vagy személynek adja meg azokat. Vagy így le fogja tölteni a számára készített, árajánlatnak álcázott vírusos fájlt stb. Mindezt milliószámra, automatikusan.

Az utánzó szélhámosság inkább a meglévő bizalmat használja ki, ez történt abban az igen hírhedt támadásban is, amiben egy bankvezető hangját és beszédmódját a banki beosztottaknak mesterségesen utánozva csaltak ki átutalásokat a támadók (Fabók, 2019). Ez ellen a videotelefonálás sem véd már meg. Ha a face2face (Thies et al., 2016) nevű amerikai, vagy a ZAO nevű kínai (Porter, 2019) valós idejű arc-cserélő szoftverek képességeiről szóló látványos videókat (Niessner, 2016) megnézzük, világossá válik, hogy akit látni és hallani vélünk, bárki lehet. A személyazonosság-lopás egyre könnyebb, már ingyenes programok és a felhasználásukhoz szükséges leírások is rendelkezésre állnak évek óta (Tran, 2017). De ha nem egyéni támadásban gondolkodnak a támadók, akkor nem csupán megváltoztatják a támadó hangját, hanem maga az MI beszélget az áldozatokkal (ahogyan a mobil asszisztensek). Például közösségi felületek rajongói oldalaira épülve, egy-egy ismert személy hangján tömegesen veszi rá őket valamire. Egyszerű közbevetéseket egy ilyen rendszer képes lereagálni, és megoldható akár az adott közösségi oldal audio-chat szolgáltatásával is, amihez még a fiókot sem kell feltörni, ha az álprofil majd azt hazudja, hogy ő a sztár igazi személyes profilja, és csak a nagy rajongókkal veszi fel a kapcsolatot.

A veszélyeket tekintve úgy tűnik a fenti módszerektől sem kell a napokban nagy horderejű offenzívára számítanunk. Inkább olyan egyedi vagy kisebb támadások várhatóak, az említett hanghamisító bankrablás mintájára, sőt, egy hamisított videofelvétel akár bűnözői alibi gyártásra is használható. Ám ezek nem nagyobb mértékűek, mint az eddigi adathalász támadások. Nagy mennyiségű személyes támadáshoz rendkívül komoly technológiai háttér szükséges, ami nem valószínű, hogy megtérül. Inkább képzelhető el egy látványos botrány, amiben egy ilyen arc-cserélt álvideót politikai lejáratásra használnak fel közösségi oldalakon.

6. Aszimmetrikus támadások

A cégek kibervédelmi szakemberei átvették az *aszimmetrikus hadviselés* kifejezést a katonai terminológiából. Abban az értelemben használják, hogy amíg a védekezésnek 100%-ban eredményesnek kell lennie rengeteg támadással szemben, addig a támadónak elegendő egyszer sikeresnek lenni számtalan próbálkozásból (Dixon - Farshchi, 2019). Vagyis több millió sikeres támadás-kivédés mit sem ér, ha egyetlen behatolás sikerül. Ha a védekező fél klasszikus védekezésnél marad, akkor az MI használatával a támadó esélyei úgy megnőnek, hogy az

aszimmetria szinte a végtelenbe fokozódik. Éppoly kivédhetetlen lesz egy megtervezett és MI alapú támadás egy klasszikus tűzfallal és víruskeresővel védett rendszer ellen, mint bőrpajzsokkal védekezni géppuskás helikopter ellen. Az arány azonban az eredeti töredékére csökkenhet, ha a védelem is MI-t használ. Az aránytalanság nem szűnik meg, elegendő egyszer bejutni – ám talán milliószoros helyett már „csak" százszoros a támadó esélye. Továbbá, bár az MI mindig új utakat talál, csak akkor lesz igazán hatásos, ha a támadó hacker is zseniálisan kreatív, hiszen az MI nem egy csodaszer, amitől egy gyenge programozóból is hatékony hacker válhat. A veszélyt összegezve úgy tűnik, hogy az aszimmetria megnövekedése a technológiai lemaradásból adódik, MI alapú védelemmel a jelenlegi szinten tartható.

7. Matematikák háborúja

Egy rendszer védelmi képességének egyik mérhető jellemzője, hogy mennyi idő alatt lehet a rendelkezésre álló technológiákkal feltörni. A gépek számítási sebességének fokozatos növekedése az addigi védelmeket rendre elégtelenné teszi. A védelem biztosítása érdekében eddig elegendő volt a titkosítási technikákat folyamatosan szigorítani és az elavultakat elvetni. Amikor azonban nagyságrendileg (ezerszeresen, akár milliószorosan) hatékonyabb technológia jelenik meg, akkor már nem védenek meg az eddigi szigorítások, mint amikor például az elvárt jelszót kicsit összetettebbre és pár karakterrel hosszabbra írták elő. Az MI terjedése ilyen nagyságrendi ugrást jelent, hiszen tanuló képességei miatt a korábbi védelmi algoritmusok matematikailag válhatnak elégtelenné. Egyes vélemények szerint a kibertérben kibontakozóban van egy MI fegyverkezési verseny, avagy a matematikák összecsapása, ahol az a kérdés, hogy melyik fél matematikája a lehető legerősebb (Da Silva, 2019). Valóban, az alkalmazott matematikai módszereken nagyban múlik, hogy a védelmet elemző kiber-támadó rendszerek milyen hatásfokkal találnak rést. De nem csupán a matematika számít. Ennek ellenvethető, hogy egy kvantum-számítógépen futtatott „brute force" (lehetőségeket sorra próbálgató) algoritmus talán gyorsabban eredményre jut, mint egy hagyományos gépen futtatott MI alapú réskereső intelligencia. (Aggasztóak e téren például Kína eredményei, akik általános célú kvantum-számítógépek mellett kifejezetten kódtörésre tervezett kvantumszámítógépet is épít, amely épp célra-tervezett volta miatt sokkal gyorsabb e téren, mint az általános célú társai). Viszont egy szuper-számítógépen futtatott MI támadás ellen mégis nehezebb védekezni (sokrétűsége miatt), ugyanis a hardver-fejlődés (akár a kvantumgépek) a védelmi oldalt ugyanúgy erősítik. Tehát mindig az alkalmazott matematikáé lesz a döntő szerep, még akkor is ha majd a kvantum-technológián sikerül neurális hálót reprezentálni.

8. Teljesen új módszerek a mesterséges intelligenciára alapozva

Eddig az MI-val erősített hagyományos támadási lehetőségeket ismertettük – viszont felmerülhetnek teljesen új módszerek is. Ilyen például a támadás célzottá tétele, az MI arc- vagy hangfelismerési képességei által. Egy másik újszerű módszerből MI-alapokra helyezett titkosítást lehet készíteni, ami az eddigieknél sokkal hatékonyabban rejti magába az információt (pl. a víruskódot). Ám az egyik leginnovatívabb irányzat a biológiai működéseken alapuló programmodellek új lehetőségeire támaszkodik. A természet évmilliók alatt megvalósított fejlesztéseiért nem kell jogdíjat fizetni, szabadon lehet belőlük ötleteket meríteni. A programozók régen is vettek innen ötleteket (maga az MI is az agy neuronhálóját modellezi). Az MI fellendülésével azonban sokkal több minden másolható, mint korábban. Már külön tudomány, a biotika foglalkozik azokkal a biológiai mechanizmusokkal, melyek programozási modellekben hasznosíthatóak. Ezek a modellek az MI kutatás egyik ágát jelentik, ugyanis ezeknél nem egy tanuló agy irányít folyamatokat, hanem „sok kis agy" közös munkája által jön létre a leghatékonyabb megoldás. Nem egy kutya agyához hasonló robotirányítás a cél, hanem olyan „virtuális lény", amit egy hangyaboly vagy darázsraj közössége testesít meg. Az ilyen

közösségek esetében a boly, raj, falka (stb). egy külön életminőséget reprezentál, és a tagok élete vagy áldozata ehhez képest eltörpül, a tagok össztevékenysége hozza létre a hatékony megoldást. Ezt használja ki kibertéri támadásokban a rajvírusok modellje. Egy másik tanulmányunkban konkrét példákon mutatjuk be ezeket és más MI alapú kibertámadásokat is (Fehér, 2019). Látható, hogy ennek a területnek a veszélyei teljesen kiszámíthatatlanok, így megítélésünk szerint a legnagyobb kockázatot jelentik.

Következtetések

Kutatásunk ezzel teljesítette fő célját, bemutattuk és értékeltük, milyen fenyegetéseket rejtenek az MI-vel kapcsolatos kibertámadások. Láthatóvá vált, hogy a veszély fennáll, és kihasználva az MI-ben rejlő lehetőségeket a kibertámadások végrehajtásakor könnyen előnybe kerülhet a támadó fél. Kérdés maradt azonban, hogy az olvasónak mit kell tenni ezzel a problémával? Felelős vezetőként vagy felhasználóként természetesen más tanulságot kell levonni - zárásul ebben szeretnénk segíteni. Az összegzéshez csoportosítsuk a fenti területeket a szerint, hogy velük kapcsolatban milyen bíztató aspektust lehet hangsúlyozni – öt féle típus jön így létre (zárójelbe a címek számát írtuk):

1. A védelmi oldalt legalább annyira segíti az MI, mint a támadót (néha jobban) (1., 2., 6.) Ezeken a területeken tehát a védelem is használja a technológiát a támadó előnye nem lesz nagyobb, mint jelenleg.
2. Ha sikeres is lesz egy támadás, az épp úgy lokális marad, csak bizonyos típusú felhasználót fog érinteni, mint az eddigi kibertámadások. (5.) Egy-egy ilyen új szélhámossági ötlet minél hatékonyabb, annál kevesebb ideig lehet alkalmazni.
3. A támadás alapja nem MI (3.,4.,7.). A megtámadott rendszer használ MI-t, tehát az ilyen támadások csak áttételesen tartoznak a témánkhoz.
4. Nem a közeljövőben várható jelentős veszély, a terület csak 5-10 év múlva válik komoly széleskörű fenyegetéssé, a világ fejlődésének ütemétől függően (5.,7.)
5. Kreatív, új módszerrel bármikor lehetséges a közeljövőben akár széleskörű MI alapú támadás is (8.) – ám ez eddig is így volt.

A Bevezető végletes hozzáállásokat taglaló fejezetében bemutattuk, hogy a közvélekedés nem is fél az MI kibertéri aktivitásától. A témát alaposan megvizsgálva arra a kissé meglepő következtetésre jutottunk, hogy a közvélekedésnek nagyjából igaza van. Egyelőre nem kell ezzel riogatni az embereket, egyelőre nincs szükség például közcélú reklámokra és ismeretterjesztő kampányokra, hogy tudjanak erről. A másik végletet, a túlzott félelmet, ami a szakma képviselői között jellemző – motivációvá kell szelídítenie minden informatikusnak. Tehát minden üzemeltető, fejlesztő vagy IT-biztonsági szakember kötelességévé válik az önképzés, illetve vezetőik tájékoztatása, tervek készítése számukra, melyek erre a fejlődésben lévő a veszélyforrásra felhívják a figyelmet. Ugyanis a jelenlegi fejlődési trendek alapján 10 éven belül eljuthatunk oda, hogy a támadó MI a mostani kibertámadásoknál nagyságrendekkel hatékonyabb lesz. Ha a védelmek nem fejlődnek a megfelelő ütemben, a támadók jelentős kárt okozhatnak elsősorban kisebb költségvetésű gazdasági szereplők tömegeinél, vagy az elmaradottabb államok informatikai rendszereiben, ez pedig világ szinten is okozhat válságos helyzetet. Tehát az állami és a vállalati vezetők számára, akik hosszabb távú tervekért is felelnek, szintén nem elhanyagolható ez a probléma. Optimális vélemény talán úgy fogalmazható meg, hogy bár pánikba esni nem szükséges az MI alapú kibertámadások lehetőségétől, de komolyan készülni kell rá. A szakújságírók és szerkesztők felelőssége, hogy ne alakuljon ki felesleges és alaptalan világ-hisztéria, amire a bevezetőben utaltunk. Jelenleg időben vagyunk, és a fenti információkból a felkészülés módja is leszűrhető: csak az MI-re alapuló védelem lesz képes felvenni a versenyt az MI alapú kibertámadásokkal szemben, tehát

az átállás elkerülhetetlen. Ebben a komoly készülésben az állami és vállalati operatív vezetők feladata, hogy biztosítsák:

1. a fejlesztésekhez szükséges a jelentős anyagi ráfordításokat;
2. az informatikai üzemeltető szakállományuk továbbképzését, hogy képessé váljanak MI alapú védelmi rendszerek tanítására és üzemeltetésére;
3. a biztonsági állomány képzését, hogy képesek legyenek a biztonsági szabályok újragondolására;
4. A fejlesztők számára olyan követelmények támasztását, ami szerint egy középfokon képzett szakember is legyen elegendő egy kibervédelmi MI napi rutin-betanítására – hiszen már most is probléma a szakemberhiány.

Amint azt más (megjelenés alatt álló) tanulmányunkban kimutattuk, technikailag kizárólag az MI alapú védelmi rendszerek alkalmazásával lehet elfogadható biztonságot megvalósítani. Ezek a rendszerek részben a fent vázolt támadó technológiákat használják védekezésre (például a mintákkal azért keresik a támadható réseket, hogy azokat ki lehessen javítani), részben a teljesen újszerű támadások detektálására alkalmazhatók, és egy indián nyomkeresőhöz hasonlóan észreveszik a behatolót akkor is, amikor a hagyományos programokat az képes becsapni. Bár mindezek által sem érhető el 100%-os védelem, de egy jelenlegihez hasonló, elfogadhatónak mondható szint tartható lehet. Zárásként, mindezek alapján megállapítható, hogy bár az MI kibertéri veszélyessége komoly kihívás, hiszen azon keresztül valós életek is fenyegetve lehetnek (például egy erőmű vagy vegyi üzem támadása, egy kórházi vagy közüzemi rendszer kikapcsolása stb. által), de léteznek az MI-nek nagyobb vagy látványosabb fenyegetései az emberiség felé. Akár oly módon, hogy egy öntanuló rendszer kicsúszik az emberi ellenőrzés alól, akár valamely autonóm fegyverrendszer által. Így egy másik kérdés nyitott marad: hogy képes lesz-e a világ társadalma arra, hogy ne következzen be katasztrófa, és egyik hiper-modern technológia (genetikai, nano, atom vagy egyéb) se váljon segítség helyett az emberiség ellenségévé? Erre tudományos módszerekkel talán nem is lehet válaszolni.

Irodalomjegyzék

1 Chartis Research (2019): The State of AI in Risk Management - Chartis Research, https://www.chartis-research.com/technology/artificial-intelligence-ai/state-ai-risk-management-10976, (megtekintve 2020. 05. 08.).

2 Da Silva, W. (2019): War of the Bots: Artificial Intelligence in Cyber Warfare, Dialogue & Discourse, https://medium.com/discourse/spy-vs-spy-cyber-warfare-gets-automated-aba60ece738c, (megtekintve 2020. 12. 07.).

3 Dickson, B. (2020): Artificial intelligence can stop IoT-based DDoS attacks in their tracks – research, The Daily Swig, https://portswigger.net/daily-swig/artificial-intelligence-can-stop-iot-based-ddos-attacks-in-their-tracks-research, (megtekintve 2021. 04. 17.).

4 Dixon, W., Farshchi J. (2019): AI is the latest weapon cybercriminals are exploiting, https://www.weforum.org/agenda/2019/09/4-ways-ai-is-changing-cybersecurity-both-in-attack-and-defense/, (megtekintve 2020. 11. 10.).

5 Durbin, S. (2020): How Criminals Use Artificial Intelligence To Fuel Cyber Attacks, Forbes, https://www.forbes.com/sites/forbesbusinesscouncil/2020/10/13/how-criminals-use-artificial-intelligence-to-fuel-cyber-attacks/?sh=5e4c8ac85012, (megtekintve 2020. 07. 01).

6 Fabók, B. (2019): Magyar céggel csaltak el tízmilliókat az első európai hanghamisításos átverésnél, G7.hu, https://g7.hu/tech/20190903/magyar-ceggel-csaltak-el-tizmilliokat-az-elso-europai-hanghamisitasos-atveresnel/, (megtekintve 2020. 06. 01.).

7 Fehér, A. T. (2019): Artificial intelligence in cyberspace 2 - realizing cyber attacks with AI, RED - American Journal Of Research Education And Development, Vol. 4, 27–41 o., https://online.pubhtml5.com/lwrb/eudn/, (megtekintve 2021.04.20.).

8 Fu, J. (2018): Security: Using AI for Evil, https://blogs.blackberry.com/en/2018/06/security-using-ai-for-evil, (megtekintve 2020. 11. 27.).

9 ITM (2020): Magyarország Mesterséges Intelligencia Stratégiája 2020-2030.

10 ITRE (2016): Industry 4.0, https://www.europarl.europa.eu/RegData/etudes/STUD/2016/570007/IPOL_STU(2016)570007_EN.pdf, (megtekintve 2020.02.14).

11 Kovács, L. (2018): A kibertér védelme, Budapest, Dialóg Campus, pp. 13, 25, 166-169, ISBN: 9786155889639.

12 McGuire, M. (2018): Into the Web of Profit, https://www.bromium.com/wp-content/uploads/2018/05/Into-the-Web-of-Profit_Bromium.pdf, (megtekintve 2020.07.20.).

13 Morgan, S. (2019): Official Annual Cybercrime Report, https://www.herjavecgroup.com/wp-content/uploads/2018/12/CV-HG-2019-Official-Annual-Cybercrime-Report.pdf, (megtekintve 2020.07.21.).

14 Muha, L., Krasznay C. (2018): Az elektronikus információs rendszerek biztonságának menedzselése, Budapest, Nemzeti Közszolgálati Egyetem, pp. 47-54, ISBN: 9786155870279.

15 Négyesi, I. (2017): A mesterséges intelligencia és a hadsereg I., Hadtudományi Szemle, X /2, http://epa.oszk.hu/02400/02463/00035/pdf/EPA02463_hadtudomanyi_szemle_2017_2_023-034.pdf, (megtekintve 2020.02.20).

16 Niessner, M. (2016): Face2Face: Real-time Face Capture and Reenactment of RGB Videos, https://www.youtube.com/watch?v=ohmajJTcpNk, (megtekintve 2021. 02. 10.).

17 Ph. Stoecklin, M. - Jiyong J. - Kirat D. (2018): DeepLocker: How AI Can Power a Stealthy New Breed of Malware, https://securityintelligence.com/deeplocker-how-ai-can-power-a-stealthy-new-breed-of-malware/, (megtekintve 2020. 11.27.).

18 Porter, J. (2019): Another convincing deepfake app goes viral prompting immediate privacy backlash, The Verge, https://www.theverge.com/2019/9/2/20844338/zao-deepfake-app-movie-tv-show-face-replace-privacy-policy-concerns, (megtekintve 2020. 11. 27.).

19 Rodriguez, M. (2019): Creating an AI Red Team to Protect Critical Infrastructure, https://www.mitre.org/publications/project-stories/creating-an-ai-red-team-to-protect-critical-infrastructure, (megtekintve 2020.11.20.).

20 Tech Wire Asia (2020): Industry 4.0 and cyber attacks: the AI-powered answer, https://techwireasia.com/2020/08/ai-ml-cybersecurity-cyber-defense-artificial-intelligence-apac-defence/, (megtekintve 2020. 12. 17.).

21 Thies, J., et al. (2016): Face2Face: Real-Time Face Capture and Reenactment of RGB Videos. Las Vegas, IEEE, 9 p., doi:10.1109/CVPR.2016.262.

22 Tran, D. (2017): Face2face — A Pix2Pix demo that mimics the facial expression of the German chancellor, Towards Data Science, https://towardsdatascience.com/face2face-a-pix2pix-demo-that-mimics-the-facial-expression-of-the-german-chancellor-b6771d65bf66, (megtekintve 2020. 11. 19.).