DOI: 10.38146/BSZ-AJIA.2025.v73.i3.pp477-494

Az automatikus biometrikus azonosság-ellenőrző rendszerek védelmének kérdései

Prezentációs támadások, megtévesztés

Questions of protection of automatic biometric identity verification systems

Presentation attacks, deception

Hazai Lászlóné
Dr., ny. nb. dandártábornok, c. egyetemi docens, kutatásért felelős főigazgatói megbízott
Nemzetbiztonsági Szakszolgálat
hazai.laszlone@nbsz.gov.hu

Absztrakt

Cél: A tanulmány célja, hogy bemutassa a biometrikus rendszerek elleni prezentációs támadások ma ismert kihívásait, ennek rendvédelmi, nemzetbiztonsági vetületeit, és felhívja a figyelmet a rendszerek alkalmazásának tervezésénél, a biometrikus személyazonosság-ellenőrzésnél a különböző szintű biztonságot jelentő megoldások, eszközök jelentőségére, lehetőségeire és ezek korlátaira, a rendszeres és célorientált kockázatelemzésre.

Módszertan: A tanulmány a prezentációs támadásokkal kapcsolatos legújabb kutatások eredményeit, a megjelent szakmai közlemények, tanulmányok, valamint teszteredmények, és irányadó nemzetközi szervezetek jelentéseinek feldolgozásával, a meglévő szabványok által nyújtott megoldások áttekintésével mutatja be a biometrikus rendszerek elleni támadásokat és von le következtetéseket.

Megállapítások: A biometrikus rendszerek mesterséges intelligencia technológiákat alkalmazó rendszerek, óriási előnyökkel és a technológia által nyújtott lehetőségekkel rendelkeznek, és persze ezzel párhuzamosan a technológiából adódó kihívásokkal, megoldandó feladatokkal. A technológia rossz célokra való használatából adódó fenyegetések, a biometrikus rendszerek megtévesztése ma valós, egyre nagyobb kihívásokat jelentő probléma, és a hamisítás, megtévesztés

A szerző a kéziratot magyar nyelven nyújtotta be. Benyújtás: 2024. 08. 14. Átdolgozás: 2024. 10. 17. Elfogadás: 2024. 10. 24.

felismerése, kivédése a feladat komplexitása okán ma még nem egyszerű. Kiterjedt kutatások folynak ma ezen a területen, de nincs általánosan használható, a rendszerekbe beépíthető megoldás vagy eszköz, alkalmazás az ilyen fenyegetések kivédésére. Egyes esetekre vannak már technikai megoldások, eszközök, de csakis kockázat alapú megközelítéssel, kockázati szintek felállításával és ennek megfelelő rendszertervezéssel és átgondolt biztonságpolitikai intézkedésekkel lehet a biometrikus rendszereket megtévesztő hamis adatközléseket, támadásokat jó eredménnyel felismerni és kivédeni, vagy a károkat mérsékelni.
Érték: Az automatikus azonosság-ellenőrzés technológia kihívásainak, a biometrikus rendszerek megtévesztése aktuális módszereinek az ismerete a védekezési módszerek, eljárások tervezésének az alapja és rendvédelmi, nemzetbiztonsági érdek is. Az elemzéssel kíván a cikk ehhez támogatást nyújtani.

Kulcsszavak: biometria, távoli biometrikus azonosítás, prezentációs támadás

Abstract

Aim: The aim of the study is to present the currently known challenges of presentation attacks against biometric systems, their law enforcement and national security aspects, and to draw attention to the importance and possibilities of solutions and devices that provide different levels of security in system design and biometric identity verification, and their limitations, to regular and target-oriented risk analysis.
Methodology: The study presents the results of the latest research on presentation attacks, by processing the professional publications, studies, test reports, reports of leading international organizations, and by reviewing the solutions provided by existing standards, it presents the attacks against biometric systems and draws conclusions.
Findings: Biometric systems are systems using artificial intelligence technologies, they have enormous advantages and the opportunities provided by the technology, and, of course, parallel to this, the challenges arising from the technology and the tasks to be solved. The threats arising from the use of technology for bad purposes, the deception of biometric systems are a real, increasingly challenging problem, and the recognition and prevention of forgery and deception is not yet easy due to the complexity of the task. Extensive research is currently being carried out in this area, but it has been established that there is no generally usable solution or tool or application that can be integrated into the systems to protect against such threats. Technical solutions and tools help, but only with a risk-based approach, setting up risk levels and corresponding system design and thoughtful security policy measures, false data communications

and attacks deceiving biometric systems can be successfully recognized and prevented or the damage mitigated.
Value: Knowledge of the challenges of automatic identity verification technology and current methods of deceiving biometric systems is the basis for planning defense methods and procedures, and is also in the interest of law enforcement and national security. The article aims to provide support for this with the analysis.

Keywords: biometrics, remote biometric identification, presentation attack

Bevezetés

Az automatikus biometrikus azonosság-ellenőrző rendszerek, különösen az arcfelismerő rendszerek fejlődése és sokcélú elterjedése különböző területeken hatalmas lendületet vett az elmúlt időszak technológiai fejlődésének köszönhetően. Ma már ezek a mesterséges intelligencia technológiák fejlett mintafelismerést, gépi tanulási algoritmusokat használnak. A 2010-es években is voltak már olyan gépi tanulási algoritmusok, amelyek segítették a nagy adathalmazok mintázatainak felismerését, azonban a 2020-as évek forrradalmi változást hoztak a generatív mesterséges intelligencia eszközök, algoritmusok fejlődésével és elterjedésével. Ezek az algoritmusok már sokkal pontosabb kimeneti adatot, eredményt adnak, jelentős mértékben javítva a biometrikus rendszerek megbízhatóságát. Ezzel párhuzamosan azonban – megfelelő szabályozás hiányában egyre jobban, és sajnálatos módon nem mindig indokolható, gyakran jogszerűtlen célok érdekében – kiszélesedett az alkalmazások köre is. Mindezzel jelentősen megnövekedett az új kihívások, megoldandó technológiai, szabványosítási, szabályozási feladatok száma, és persze ahogy ez lenni szokott, a rosszindulatú felhasználások, visszaélések, a veszélyek mértéke is.

Régóta lehetséges digitálisan megtévesztő képeket, videókat vagy hanganyagokat készíteni. A mesterséges intelligencia technológiák által ez még könnyebbé vált, és ennek a következménye, hogy az ember biológiai jellemzőire épülő automatikus azonosság-ellenőrző rendszerek biztonsági fenyegetései is egyre nagyobb aggodalomra adnak okot.

A biometrikus azonosság-ellenőrző rendszerekben az identitás-ellenőrzés számos alkalmazási területén (intézménybe, helységbe, számítógépbe, mobil eszközökbe történő belepés engedélyezésénél, digitális szolgáltatások igénybevételénél, elektronikus pénzügyi tranzakciók lebonyolításánál, ellenőrzésénél, a különböző céllal történő információgyűjtésnél, adatgyűjtésnél stb.) a manipulált arcképek automatikus feldolgozása téves döntéseket eredményez. Ezért

a manipulált arcképek valósnak megfelelő elfogadása súlyos károkhoz, következményekhez vezethet.

Kockázatok, kihívások

A biometrikus rendszerek és ezen belül az arcfelismerő rendszerek piaca a Grand View Research (2020) piackutató cég szerint 2022 és 2030 között várhatóan jelentősen növekszik. Értékelésük szerint ennek oka egyrészt a rendészeti, határvédelmi személyazonosság-ellenőrzést végző biometrikus alkalmazások számának növekedése, másrészt az ilyen alkalmazások okostelefonokba történő széles körű integrálása, a jelszó nélküli identitáshitelesítés elterjedése (lásd a telefon feloldása, a mobilalkalmazásokba való bejelentkezés és a pénzügyi tranzakciók ellenőrzése stb.).

Arcfelismerő algoritmusokat használnak a felhasználók biometrikus hitelesítésére, távoli biometrikus személyazonosításra, az identitások adatbázisokban történő megkettőzésének megszüntetésére, az adatbázisban lévő identitások biometrikus összehasonlítására. A biometrikus jellemzőinket használó, arcfelismerő, ujjlenyomat- vagy íriszazonosító technológiával ellátott okostelefonok és számítógépek már a jelen eszközei.

A biometrikus azonosság-ellenőrzés technológiája – a számítástechnikai kapacitás, a mesterséges intelligencia technológia gyors fejlődésének köszönhetően – egyre pontosabbá, biztonságosabbá vált. Azonban csupán a technológia alkalmazása nem jelent teljes biztonságot, mert nem nyújt teljes bizonyosságot mindig, minden esetben a felhasználó személyét illetően a biometrikus jellemzők hamisítási lehetőségei, a biometrikus rendszerek megtévesztésének lehetőségei, de a technológia esetleges sérülékenységei, statisztikai hibái miatt sem.

A digitális szolgáltatások igénybevétele egyre fontosabb az életünkben, azonban ha a műveletek során a felek kilétében nem lehet megbízni, az komoly probléma. Tény, hogy a biometrikus rendszerek algoritmusainak teljesítménye és pontossága folyamatosan javul, és a felmerülő kihívások kezelésére is újabb és újabb fejlesztésekkel reagál a kutatók közössége és az ipar – lásd arcfelismerés maszkban a COVID időszakban (Ngan et al., 2022) –, azonban még nagyon sok feladat vár megoldásra, így többek között a hamisítások megakadályozásának, időben történő felismerésének kérdéskörében is.

A lehetséges fenyegetések sikeres elhárítása érdekében fontos megismernünk a lehetséges támadási módszereket, eszközöket.

Az arcfelismerő rendszerek esetében, a bemenő adatokat érintően az arckép-manipuláció a nagy kihívás. Ennek különböző indoka lehet, és módszereit

tekintve is nagyon sokféle megoldás fordulhat elő. Lehet az ok esztétikai, szépészeti. Ebben az esetben ugyan nem rosszindulatú a művelet, de fontos tudni, hogy ez a korrekció is ronthatja az arcfelismerő rendszer teljesítményét, a megfeleltetés eredményét. Lehet azonban egy rosszindulatú művelet, amikor valaki szándékosan megtéveszti a biometrikus azonosság-ellenőrző rendszert úgy, hogy például egy személy más személynek adja ki magát, vagy az illető el akarja kerülni azt, hogy a rendszer felismerje. A biometrikus azonosság-ellenőrzésnél, mind az azonosság megerősítése (hitelesítés), mind az azonosság megállapítása (felismerés) esetében ez utóbbi már hamisítás és napjainkban is széleskörűen kutatott téma.

A technológiai fejlődéssel egyre több olyan új technika, eszköz jelenik meg, amelyek megkönnyítik az ilyen rossz szándékú műveletek egyre jobb minőségű végrehajtását (Zhao et al., 2021). Ezt segíti az is, hogy a módszerek, eszközök, a manipulációs algoritmusok és a használatukhoz szükséges útmutatók széles választéka ingyenes webes és mobilalkalmazásokban mindenki számára hozzáférhető. A manipuláció megvalósulhat számítógépes grafikán alapuló módszerekkel (például Face2Face) is, de mélyhamisítás (deepfake) eszközökkel (például Face App4, DeepfakeApp ZAO) nagyságrendekkel jobb eredmény érhető el, és ahogy a technológia fejlődik egyre jobb, valósághűbb, nyilvánosan elérhető megoldások jelennek meg. Tehát nő a kockázat is, ami megköveteli az ellenintézkedéseket.

A *European Union Agency for Law Enforcement Cooperation* jelentés (Europol, 2022) foglalkozik a mélyhamisítás technológia szerepével és hatásával a különféle bűncselekmények alakulására. A jelentés megállapítja, hogy „a bűnözők az új technológiák korai alkalmazói". Felhívja a figyelmet arra, hogy a mélyhamisítás technológiák segítik az okmányhamisítást és az online személyazonosság meghamisítását. Továbbá figyelmeztet arra a tendenciára, hogy ezen a területen is jellemző, és a prognózis szerint gyors fejlődésnek indult egyfajta bűnözésszolgáltatás (CaaS). Az ilyen típusú szolgáltatás hozzáférést árul új technológiájú eszközökhöz, módszerekhez, és ismeretekkel, gyakorlati útmutatással segíti a bűncselekmények elkövetőit.

A jelentés megállapítja azt is, hogy a magas színvonalú okmányvédelmi megoldásoknak köszönhetően ma egy útlevelet, egy jól megtervezett személyazonosító okmányt egyre nehezebb hamisítani, de a digitálisan manipulált képek felhasználása új helyzetet, megközelítést jelent. Az új technológiák, a mélyhamisítási módszerek segítik a rossz szándékú elkövetőket, a szervezett bűnözői csoportokat abban, hogy egy nem valós, manipulált arckép, egy meghamisított, de eredeti okmány átmenjen a személyazonosság-ellenőrzéseken, beleértve az automatizált arcfelismerő ellenőrzéseket is.

Ugyan a megtévesztő képek egyes esetekben tartalmazhatnak olyan hiányosságokat, amelyek alapos humán vizsgálattal még észrevehetőek (például inkonzisztencia a képen, egyes képminőségi hibák megjelenése, a szemben fényvisszaverődés, pislogás hiánya a videón stb.), de a jelentés megállapítja, hogy nagyobb hatékonysággal lehet a valódiságot automatikus – mesterséges intelligencia – módszerekkel vizsgálni.

Mik ezek a módszerek? Akhtara és munkatársai (2019) a cikkükben áttekintést nyújtanak az arckép-manipuláció generálásának és a manipuláció felismerésének technikáiról. Mindkét terület módszerei folyamatosan fejlődnek, ebből is következik, hogy a technológiai megoldások ma még kutatott témák, és sok még a megoldandó kérdés. A kihívásokat jellemzi a hivatkozott szerzők néhány kiemelt, részben ma is aktuális megállapítása, miszerint 1) a manipuláció felismerésére fejlesztett detektorok csak bizonyos típusú manipulációk felismerésére használhatók jó teljesítménnyel; 2) a legtöbb felismerési technika alkalmatlan mobilalkalmazásokhoz; 3) a felismerés technikai fejlesztéséhez, a tanításhoz nincs elegendő jó minőségű adatkészlet.

Fizikai és digitális prezentációs támadások

A biometrikus arcfelismerő rendszerek elleni leggyakoribb manipulációs technika a prezentációs támadás. A prezentációs támadások alapvetően a biometrikus rendszer bemenő adat szintjén valósulnak meg, és leginkább olyan alkalmazásokban, amelyek esetében nincs humánfelügyelet. Az elkövető célja ezekben az esetekben a sajáttól eltérő identitás igazolása, vagy a felismerés elkerülésével az azonosság-ellenőrzés megakadályozása. A prezentációs támadások során az arcfelismerő rendszerek sérülékenységeit, sebezhetőségeit használják ki az elkövetők.

Az arcfelismerő rendszerek esetében a rossz szándékú támadásoknak (csalás vagy személyazonosság-lopás, a biometrikus rendszerek megtévesztése), a manipulációs technikáknak, módszereknek igen sokoldalúan feldolgozott irodalma van. Ezen munkák a prezentációs támadásokat különböző csoportokba rendezik (Akhtara et al., 2019; Garrido et al., 2014; Hernandez-Ortega et al., 2019; Qin et al., 2021; Yu et al., 2023).

Rathgeb és munkatársai (2024) például az arcmanipulációnak hat, ma is a kutatások fókuszában álló típusát, módszereit elemzi: 1) teljes arcszintézis, 2) identitáscsere, 3) morfizálás, 4) attribútum-manipuláció (az arckép egyes tulajdonságainak módosítása), 5) arckifejezés manipuláció („videoreplay” támadás) és 6) videoarc és -hang manipulációk.

A prezentációs támadások, módszerek palettája tehát rendkívül széles, vannak a biometrikus rendszereket megtévesztő, úgymond „egyszerű” fizikai technikák, eszközök és vannak digitális megoldások, fejlett, mesterséges intelligencia technológiájú támadástípusok. Ennek megfelelően lehet különböző csoportokba sorolni ezeket a módszereket, így vannak:

- Fizikai módszerek:
  - az identitás elrejtése különböző módszerekkel (szemüveg, smink, paróka stb.),
  - a valós identitás helyett egy megtévesztő, kinyomtatott idegen kép felmutatása, 3D maszk alkalmazása.
  - Külön csoportot képez az úgynevezett „videoreplay” támadás, amikor az identitást igazoló valós videófelvétel helyett egy manipulált videó bemutatására kerül sor.
- A digitális támadás – ezek az úgynevezett digitális „injekciós” támadások –, amely egy manipulált digitális kép vagy videó elektronikus injektálását jelenti a biometrikus rendszerbe, továbbá ebbe a körbe sorolt támadás típusok még a mélyhamisítás vagy deepfake támadás és a morfizálás (a képek szintetizálása) is.

Megjegyzendő, hogy míg a hagyományos pezentációs támadások az adatrögzítés szintjén (a kameránál), az injekciós támadások már nem itt valósulnak meg, hanem a következő szinten, a biometrikus felismerést végrehajtó alrendszer (az összehasonlítás) szintjén, és ez már lehetőséget teremthet a rendszerbe beépített automatikus támadásészlelés elkerülésére is.

Firc és munkatársai (2023) egy friss publikációjukban az előzőeket kiegészítve nyújtanak még részletesebb áttekintést a ma lehetséges támadástípusokról. A hivatkozott szerzők szerint ezek a prezentációs támadási módszerek a következők lehetnek:

- Arcképcsere, másik ember arcképének a bemutatása a saját helyett.
- Arckifejezés animáció, a lecserélt arckép „életre keltése”, vagyis például arckifejezés, mozgás jellemzőkkel való kiegészítése (lásd még Garrido et al., 2014).
- Arcmanipuláció, ez már az eredeti arckép vagy videó digitális manipulációja, ami jelenti az arc legkülönbözőbb egyedi vagy egyéb fizikai, érzelmi jellemzőinek digitális megváltoztatását, de például manipulációs módszer az is, amikor speciális perturbációkat (azaz zavarásokat) adnak az eredeti képhez, mert ez is megtéveszthető a biometrikus rendszereket. Akhtara és munkatársai (2019) szerint az ilyen hamis képek, videók létrehozásához már fejlett neurális hálózatot is használtak a jó minőség érdekében: egy GAN

(Generatív Adversarial Network – versengő neurális hálózat) generatív hálózatot és egy FaceSwap technikával rendelkező diszkriminatív hálózatot.

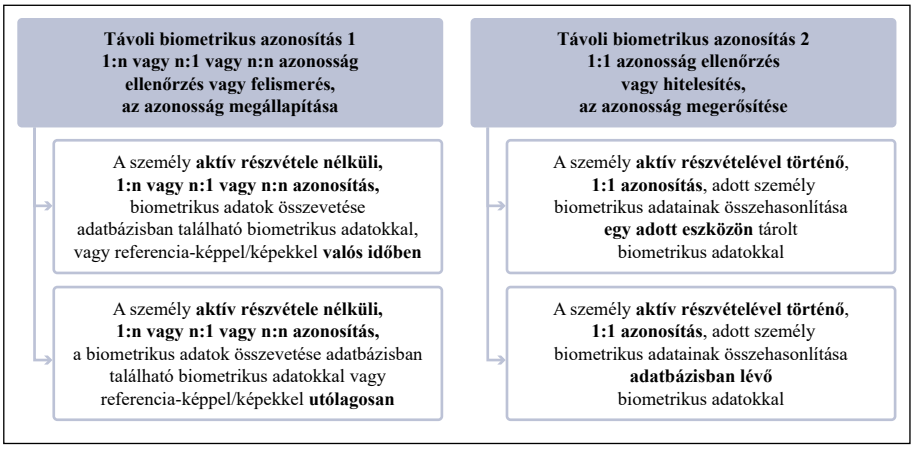
- Arcszintézis, nem létező arcok digitális úton történő előállítása magas szintű attribútumok alapján. Az arcszintézis módszerek számossága is igen nagy, ennek az oka, hogy ezeket használják virtuális karakterek előállítására is (lásd szórakoztatóipar, virtuális asszisztensek), és gyakran a biometrikus rendszerek, jelen esetben az arcfelismerő rendszerek betanítására a valós adatok helyett. Ez a módszer is alkalmas visszaélésre, hamis személyazonosság létrehozására, az identitás elrejtésére.
- Az arcmorfizálás egy olyan különleges képmanipulációs technika, ahol két vagy több személy arcképjellemzőinek statisztikai átlagát képezve egyetlen új, „szintetikus" arcképet hoznak létre, amely ilymódon az összes közreműködő (kettő vagy több) személyre megtévesztésig hasonlít. A morfizált képpel tehát több közreműködő is megtévesztheti a humán ellenőrzést (Kramer et al., 2019), de még az automatikus arcfelismerő rendszereket is. A morfizált kép létrehozása ma már különféle mesterséges intelligencia technikákkal valósítható meg, az egyszerűbb képmanipulációs módszerektől a legújabb GAN alapú generatív technikáig.

  A bűnügyi célú arcmorfizálás ismert módszerei: eredeti, lopott személyazonosító okmány meghamisítása, vagy hamis adatokkal egy új okmány igénylése, és a manipulált képet tartalmazó okmánynak a felhasználása. Több eset, sikeres megtévesztés, csalás is nyilvánosságra került már, és nem kérdés, hogy ez a módszer nagy problémát jelent bármilyen jogosultság-, hitelességigazolás esetében. Sajnálatos tény, hogy a morfizálás megvalósítására számos eszköz a neten szabadon elérhető (Venkatesh et al., 2021).

A kockázatok csökkentése, a prezentációs támadások észlelése

A manipulált arckép a távoli személyazonosság-ellenőrzést (módszereit lásd 1. számú ábra), az azonosság-ellenőrzés minden formáját, különösen a személyazonosság hitelességét megkérdőjelező potenciális fenyegetések, csalások, visszaélések eszköze lehet. De a biometrikus rendszerek iránti bizalom elvesztéséhez is vezethetnek, mivel a hamis adatok „szennyezik" az adatbázist és megtévesztik a rendszereket, amelyek így rossz döntéseket hoznak. Tehát ez a rendszerek megbízhatóságát is befolyásoló tényező.

1. számú ábra
A távoli biometrikus azonosítás lehetséges esetei (az EU 2024/1689 rendelete alapján)



Forrás. A szerző saját szerkesztése.

Az elmúlt években a kutatók és az ipar újabb és újabb feljesztésekkel az arcfelismerő rendszerek elleni prezentációs támadások észlelését jelző megoldásokat (úgynevezett PAD módszereket, modulokat) integráltak a különböző biometrikus rendszerekbe. Az automatizált prezentációs támadásészlelési módszerekkel bizonyos támadások kockázatai csökkenthetők és javítható a rendszerek biztonsága, akár felügyelt, akár nem felügyelt adatrögzítés történik. Ezek a PAD módszerek, modulok is mesterséges intelligenciát, gépi tanulást használnak a képek, videók képi jellemzőinek elemzésére. A The European Union Agency for Cybersecurity megállapítása (ENISA, 2022), hogy a videó értelemszerűen több adatot biztosít a PAD módszerek részére, ami csökkentheti a személyazonosság csalás lehetőségét.

Az arcfelismerés esetében ezek a támadásészlelési módszerek elsősorban olyan, a valós és a támadó prezentációk közötti lényeges minőségi különbségeket vizsgálják, mint például a fiziológiai jellemzők alapján az életszerűséget, a különböző mozgásmintákat és a színeltéréseket (Qin et al., 2021).

A PAD módszerek fejlesztésének támogatására született meg az ISO/IEC 30107 szabvány (ami egy szabványsorozat). Az ISO/IEC 30107-1 definiálja a prezentációs támadásokat, és az előállítás eszköze, valamint az emberi jellemzők (megváltozott vagy megváltoztatott, vagy nem megfelelő minőségű) szerint kategóriákat állít fel. Az ISO/IEC 30107-2 adatformátumokat határoz meg. Az ISO/IEC 30107-3 pedig elveket állít fel, és módszereket határoz meg a különböző PAD mechanizmusok teljesítményértékeléséhez. A mobileszközök

jellemzői és sokfélesége tette szükségszerűvé egy külön szabvány kidolgozását, ezért az ISO/IEC 30107-4 követelményeket ír elő a zárt rendszerként működő, helyi (1:1) biometrikus felismeréssel rendelkező, biometrikus mintákat rögzítő, összehasonlító és biometrikus sablonokat tároló mobileszközökhöz illeszthető PAD mechanizmusok teljesítményének értékelésére.

A szabványsorozat azonban a prezentációs támadások csak bizonyos típusainak észlelésére szorítkozik. Nem nyújt megoldást az olyan támadások észleléséhez, amelyek a mesterséges intelligencia technológia gyors ütemben létrejött újabb fejlesztései, és ahol a két vagy több biometrikus adatalanynak megfeleltetése érdekében manipulált biometrikus bemeneti (morfizált) mintákat injektálnak a rendszerekbe.

Az automatikus támadásészlelési módszerek teljesítményértékelését illetően az ISO/IEC 30107-3 szabvány szerint is egy PAD modul szerepe, a célnak megfelelő alkalmassága a támadás típusa és kivitelezése mellett rendkívül sok más tényezőtől függ, így többek között magától a biometrikus azonosság-ellenőrző rendszertől, amelybe a PAD modult integrálják. De fontos tudni azt is, hogy minden automatikus támadásészlelési módszernél is fennáll a hibás döntés lehetősége. [A hamis pozitív jelzések (FAR) tévesen minősíthetik a valós bemeneti adatokat prezentációs támadásnak, lerontva a rendszer hatékonyságát, illetve a hamis negatív jelzések (FRR) esetében a rendszer a prezentációs támadásokat valós bemeneti adatként rögzítheti, lerontva a rendszer biztonságát.]

A támadás-detektáló módszerek értékelése

A rossz szándékkal előállított képi deepfake adatok észlelése, és a biometrikus rendszerek megvédése az ilyen típusú támadásokkal szemben társadalmi szinten is egyre nagyobb érdek. Ezért igen nagyszámú kutatás foglalkozik a különböző támadás-detektáló módszerek kidolgozásával, fejlesztésével, és a detektorok hatékonyságának értékelésével, a kidolgozott módszerek általánosíthatóságának kérdésével. Az általánosíthatóságot illetően a kutatók megállapítása ma, hogy a különböző módszerek és hatékonyságuk összehasonlítása (Dong et al., 2023; Yan et al., 2023) még mindig téves, félreértelmezhető eredményekhez vezet, annak okán, hogy az egyes fejlesztések bemeneti adatai különbözőek, inkonzisztensek, de különbségek vannak a vizsgálati paraméterek, az értékelési stratégiák és a képminőség összehasonlításához a mérőszámok kiszámításának módszereit illetően is. Ezen problémák kiküszöbölésére dolgoztak ki a kutatók a legkorszerűbbnek tartott detektáló rendszerekre egy keretszabályt. De az általánosíthatóság kérdése továbbra is a kutatások fókuszában van.

A National Institute of Standards and Technology (Nemzeti Szabványügyi és Technológiai Intézet – NIST) az USA egyik legrégebbi fizikai tudományos laboratóriuma. Sok éve végzi azokat az arcfelismeréssel összefüggő technológiai teszteket, amelyek során – általuk meghatározott szabályok alapján, tesztesetekkel, szabványos feltételek mellett – értékelik, elemezik a fejlesztők algoritmusainak képességeit, hiteles információkat szolgáltatva arról, hogy a modern képfeldolgozó szoftverek milyen hatékonyan hajtanak végre egy feladatot. A NIST a korábban elindított nemzetközi arcfelismerő (FRVT) tesztprogramja – az 1:1, és 1:n algoritmusok pontosságának, teljesítményének vizsgálata, értékelése – ma már kétirányú. Az egyik irány az arcfelismerő technológia értékelése (FRTE) és a másik az arcelemző (a képeket jellemző, minőségelemző) technológia értékelése (FATE).

A legújabban lefolytatott *Arcelemzési technológia értékelése (FATE) 10* című elemzés (Ngan et al., 2023) a passzív, szoftveralapú PAD algoritmusok teljesítményével foglalkozik. Ez a jelentés dokumentálja több szoftver alapú PAD algoritmus teszt teljesítését, pontosságát. De ezek a tesztek is csak a prezentációs támadás típusok egy részterületét fedik le, mivel a vizsgálatok a fizikai prezentációs támadások azon teszteseteire vonatkoznak, ami alatt a különböző 2D képpel és a videofelvétellel végrehajtott támadások értendők. A jelentés 45 fejlesztő 82 PAD prototípus algoritmusának tesztadatait elemzi.

A jelentés fontos általános megállapítása ezeknél a támadástípusoknál, hogy a támadásészlelés pontosságát nagymértékű változékonyság jellemzi, az algoritmusoktól és a felhasználási esetektől (más személyi adataival való visszaélés/a felismerés akadályozása), valamint a prezentációs támadás kivitelezésétől függően. Egyes prezenzációs támadások esetében több PAD algoritmus is jól teljesített, ugyanakkor más támadástípusok esetében ugyanezek az algoritmusok magas támadásészlelési hibaarányt adtak.

A teszteredményekből a NIST megállapítása többek között az is, hogy az egyes PAD módszerek teljesítménytesztjei nem összehasonlíthatók. Amíg a biometrikus rendszerek teljesítménytesztje során megvalósíthatók a standard körülmények, a PAD teszteknél nem, mivel a prezentációs támadás típusától, az alkalmazott eszközöktől függően nagyon sokféle teszt megközelítés lehetséges. A PAD teszteredményeket a támadás kivitelezésének módja, minősége, de még a tesztelő képessége, felkészültsége is befolyásolja. A jelentés megállapítása, hogy a PAD tesztek során kapott teljesítménymutatók, hibaértékek kizárólag csak az adott, vizsgált PAD mechanizmusra jellemzőek.

Az NIST 2018 óta végez morfizálással végrehajtott támadástípusokkal kapcsolatos teszteket is. A tesztek során különböző morfizálási módszerrel létrehozott adatkészletet használnak. A legfrissebb NIST MORPH tesztek (Ngan

et al., 2024) során vizsgálták a különböző támadásdetektáló algoritmusok teljesítését, a 2D és 3D képek minőségének, felbontásának hatását a detektálás eredményességére. A jelentés szerint, ha csak egy képről kellett az algoritmusoknak eldönteni azt, hogy a kép valódi vagy morfizált-e, többnyire magas tévesztési eredményeket kaptak, és ebben az esetben is az eredmény algoritmusfüggő volt, ilymódon nem általánosítható. Megállapították továbbá azt is, hogy egy kinyomtatott digitális kép, majd ennek az újradigitalizálása az egyik legnagyobb kihívás a morfizálás detektálása szempontjából. A különböző módon előállított, különböző minőségű morfizált képek detektálásánál továbbra is jellemzően magas volt a hamis detektálási arány. Jobb eredmények születtek, vagyis csökkent a téves detektálás abban az esetben, ha két olyan képet vizsgált az algoritmus, ahol az egyik egy feltételezett morfizált kép volt, a második pedig egy valós élőkép az érintett személyek egyikéről (például élő felvétel egy határellenőrzési pontnál). A szoftvernek ebben az esetben is döntést kellett hoznia arról, hogy a kép morfizált-e, és megbízhatósági pontszámot kellett adnia a döntéséhez. Ezt a jobb eredménynek tartott pontszámot azonban lerontotta, vagyis ismét nőtt a tévesztés mértéke, ha jelentősen hosszú idő telt el (öregedés) a két képfelvétel között.

Ezek az NIST tesztek azt mutatták, hogy az alacsony minőségű, kevésbbé sikeres morfizálás automatikus azonosítása viszonylag jó eredménnyel megvalósítható, mert 100-ból „csak" 15 morfizált képet hagyott figyelmen kívül egy jónak tartott algoritmus. De nyilvánvaló, hogy ez az eredmény valós használati esetekben (útlevél, úti okmány ellenőrzés, eGate átlépés) azért nem igazán megnyugtató. A jó minőségű morfizálás azonosításánál ettől rosszabb volt az eredmény, mert a legjobban teljesítő algoritmusnál is kaptak olyan magas hibarányt, ami azt jelenti, hogy a morfizált képek 88%-a átmenne az ellenőrzésen, úgy mintha az eredeti lenne, és ez már valós esetekben nem elfogadható kockázat.

Tehát a támadási módszerek sokszínűsége, a morfizálással előállított „termékek" minőségének széles skálája is oka annak, hogy jelenleg nincs szabványos módszer, általánosan használható támadásészlelési mechanizmus, olyan – a biometrikus arcfelismerő rendszerbe integrálható – algoritmus, szoftver, eszköz, amely minden szempontot (használati esetet, támadási módszert) figyelembe tud venni, és általánosan véd a rosszindulatú támadási műveletek ellen.

A személyazonosság-lopás elkerülése érdekében a PAD módszerek beépítése a biometrikus rendszerekbe ugyanakkor ajánlott. Azonban tekintettel az előzőekre, a fejlett biometrikus azonosság-ellenőrzési technológiák pontossága, képessége ellenére az irányadó szervezetek (ENISA, ISO/IEC, ICAO) nem javasolják a biometrikus adatok – különösen az arc jellemzőinek – egyedüli, önálló, hitelesítési tényezőként való használatát.

A prezentációs támadások felismerése, a kockázatok csökkentése

Arra az Europol 2022. évi jelentése is rávilágít, hogy a deepfake technológiákkal okozott személyazonosság-lopások, -hamisítások reális kockázata igen magas, és az ilyen fenyegetések hatékony kezelése a társadalom minden szegmensének, szereplőjének az érdeke (Europol, 2022).

Ennek megfelelően a jelentés is rögzíti, hogy a kockázatok kezelése érdekében a biometrikus rendészeti jellegű alkalmazások, az online szolgáltatások vonatkozásában alkalmazni kell megelőzést szolgáló különböző PAD technológiákat. Mivel az előzőek alapján egyértelmű, hogy nincs minden (jelenleg ismert és jövőbeli) támadástípus felismeréséhez általánosan alkalmazható módszer, eszköz, ezért a jelentés megállapítása, hogy a biometrikus azonosság-ellenőrzési rendszerek alkalmazásának gyakorlatához ki kell dolgozni a rendvédelemnek, a különböző szolgáltatóknak az érdekeiknek megfelelő, a biztonságot szavatoló, az emberek jogait figyelembevevő védekezési metodikát, biztonságpolitikát.

Az elemzések, értékelések alapján tehát a megállapítás, hogy biztonsági szinteket kell felállítani és ezekhez keretszabályokat kell meghatározni, továbbá folyamatosan fejleszteni kell az ellenőrzés módszereit.

Az Európai Unió Kiberbiztonsági Ügynökségének (ENISA) tanulmánya (ENISA, 2022) a prezentációs támadásokat elemezve így fogalmaz: *„A lehetséges fenyegetések elleni küzdelemhez először tisztában kell lennünk velük, valamint meg kell értenünk a támadók rendelkezésére álló eszközöket és módszereket. Nemcsak a biometrikus adatok ellenőrzésének védelme fontos; az egész folyamatot pontról pontra biztosítani kell. Elemeznünk kell a fenyegetéseket, kockázatalapú megközelítést kell alkalmaznunk az ellenintézkedések prioritása érdekében, és mindig figyelemmel kell lennünk az új trendekre, hogy megelőzzük a támadókat.”*

Az ENISA pezentációs támadásokkal foglalkozó ez évi jelentése (ENISA, 2024) megállapítja, hogy a távoli személyazonosság-igazolás – a hitelesítésnek e módszere – a digitális szolgáltatás kritikus eleme. A prezentációs támadások a technológia fejlődését gyorsan követő fenyegetést jelentenek. Az elemzés szerint a prezentációs támadások néhány korábbi módszere, feltehetően az egyszerűség, a gyorsabb kivitelezhetőség okán ma még népszerű, de a legnagyobb veszélyt a mélyhamisításos és az injekciós támadások jelentik, mert ezek nehezen felismerhetőek és nehezen kivédhetőek. Annak ellenére, hogy ez utóbbi támadástípusok nagyobb felkészültséget igényelnek, az ilyen támadások száma 2022 év óta növekvő tendenciát mutat, és a végrehajtás technikája is egyre kifinomultabb.

Ez az ENISA jelentés foglalja össze az arcfelismerésen alapuló, biometrikus azonosság-ellenőrzési rendszerek alkalmazásának a biztonság és a bizonyosság

érdekében betartandó és elvárt, a pezentációs támadások feismerésével és elkerülésével kapcsolatban jelenleg számbavehető intézkedéseket. Ezek az ENISA ajánlása szerint következők lehetnek.

1) Konkrét, az adatfelvétellel kapcsolatos technikai feltételek, előírások. Például:
   - megfelelő fényviszonyok biztosításának előírása az adatrögzítésnél,
   - minimum multimédiás előírások (technikai követelmények, fotó és videó minőségi kritériumok) meghatározása,
   - dedikált kliensoldali távoli személyazonosság-ellenőrzési alkalmazás használata (a webalkalmazás helyett).
2) A prezentációs támadások észlelésére PAD módszerek, modulok alkalmazása.
3) IAD (Injection Attack Detection – injekciós támadásdetektálás) módszerek alkalmazása.
   A szabványosított támadásészlelési PAD módszerek jelenlegi hiányosságai, illetve kikerülésének lehetősége miatt szükséges az IAD módszerek alkalmazása is. Ezek olyan módszerek, amelyek a biometrikus rendszer legkülönbözőbb szintjein valósulhatnak meg. Lehetnek megelőző (ilyen például a kriptográfiai képtanúsítvány, az adatrögzítő kamera azonosítása vagy a szteganográfiai vízjel alkalmazása) és felderítő (ilyen például a munkamenet-, metaadatelemzés egy virtuális kamera vagy eszköz beiktatásának észlelése céljából) jellegű.
4) Bemutatott szeményazonosító okmány eredetiségének az ellenőrzése (az okmánybiztonsági és az elektronikus biztonsági elemek meglétének, eredetiségének az ellenőrzése).
5) Különböző folyamat-kontroll szabályok, módszerek alkalmazása.
   Az ENISA szerint ilyenek:
   - az úgynevezett kihívás-válasz protokollok, amely lehet egyben egy multimodális biometrikus ellenőrzés is (az arckép és a hang alapján a beszélő azonosítása),
   - ide sorolható a humánoperátor ellenőrzés, valamint
   - a megbízható, hiteles azonosítási források, adatbázisok használata.
6) A biometrikus rendszerek alkalmazásának, üzemeltetésének folyamatába épített biztonsági, kiberbiztonsági szabályok felállítása, betartásának ellenőrzése.

Következtetések

Az automatikus biometrikus azonosság-ellenőrző rendszerek támadásokkal szembeni védelme az alkalmazási területek, a lehetséges prezentációs támadások sokszínűsége miatt csak komplex, célirányos, kockázatalapú, optimálisan megtervezett biztonsági intézkedések mellett lehet elvárt hatékonyságú. A 1183/2024/EU rendelet – az eIDAS ökoszisztémák tervezése – külön aktualitást ad az automatikus biometrikus azonosság-ellenőrző rendszerek támadásokkal szembeni védelmi kérdéseinek, mert a technológia fejlődése, a technológiai nyitottság, valamint a szabályozás el-, illetve lemaradása könnyebbé tette a rossz szándékú műveletek megvalósításának lehetőségét. Vannak jól és vannak nehezen felismerhető és igen nehezen kezelhető fenyegetések, támadási módszerek. Az ENISA (ENISA, 2022) által javasolt, a prezentációs támadások kockázatait mérséklő intézkedési megoldások és eszközök alapján mondható, hogy sok, különböző hatékonyságú, és különböző biztonsági szintet képviselő módszer és technika áll ma már rendelkezésre, amelyek lehetőséget adhatnak a prezentációs támadások felderítésére, a kockázatok mérséklésére.

Nagy intenzitással folynak a kutatások a megbízható automatikus biometrikus személyazonosítás, a prezentációs támadásészlelés módszereinek kérdéskörében. Azonban az ilyen támadásokat csak a biometrikus rendszerekre és a kapcsolódó eszközökre vonatkozó – az előzetes kockázatelemzésnek, a különböző (jogosultsági, használati) szempontok alapján meghatározott biztonsági besorolásnak megfelelő – konkrét technikai követelményekkel, az optimálisan megválasztott beépített eszközökkel és a különböző eljárási intézkedések, szabályok együttes előírásával lehet csak kezelni. Fontos javaslata az ENISA-nak a rendszeres fenyegetettségi kockázatelemzés, mert a technológia gyors fejlődése a védelemi megoldások és a támadási módszerek szempontjából is folyamatos figyelmet követel.

Az eIDAS rendelet szerint az elektronikus azonosítási rendszerhez meg kell határozni az elektronikus azonosító eszközöknek tulajdonított „alacsony”, „jelentős” és/vagy „magas” biztonsági szintet. A jogszabály szerint ez a biztonsági szintbe történő besorolás attól függ, hogy az elektronikus azonosító eszköz milyen szintű megbízhatósággal igazolja egy személy személyazonosságát. Vagyis milyen megbízhatóságú technológia, rendszer, eljárás, intézkedés van egy személy hitelességét igazoló elektronikus azonosító eszköz mögött. Ha a hitelesség megállapítása biometrikus azonosság-ellenőrzést jelent, akkor a rendszernek a megbízhatóságot illetően a rendeletnek is meg kell felelni.

A megbízhatóságot pedig a biometrikus személyazonosság-ellenőrző rendszerekbe integrált, a biztonsági szintekhez illesztett mechanizmusokra és eszközökre

vonatkozó konkrét technikai specifikációkkal, szabványokkal és intézkedésekkel, eljárási szabályokkal, illetve folyamatos ellenőrzésekkel, tesztekkel lehet alátámasztani.

A biometrikus rendszerek megtévesztésének kockázatai, a lehetséges támadási módszerek variációinak sokszínűsége, de a biometrikus rendszerek sérülékenységei, a mesterséges intelligencia technológia ma még meglévő anomáliái, hibái miatt is fontos a biometrikus rendszerek alkalmazásának megfontolt kockázatalapú megtervezése, illetve megvalósítása.

Felhasznált irodalom

Akhtara, Z., Dasgupta, D., & Banerjee, B. (2019). Face authenticity: An overview of face manipulation generation, detection and recognition. *Proceedings of the International Conference on Communication and Information Processing (ICCIP-2019),* pp. 1–9. Elsevier-SSRN. https://www.researchgate.net/publication/333984240

Dong, F., Zou, X., Wang, J., & Liu, X. (2023). Contrastive learning-based general Deepfake detection with multi-scale RGB frequency clues. *Journal of King Saud University – Computer and Information Sciences, 35*(4), pp. 90–99. https://doi.org/10.1016/j.jksuci.2023.03.005

ENISA. (2022). Remote identity proofing: Attacks & countermeasures. In V. Paggio, E. Nikolouzou, & M. Dekker (Eds.), *ENISA report.* Publications Office of the European Union. ISBN: 978-92-9204-549-4. https://www.facetec.com/wp-content/uploads/2022/01/ENISA-Report-Remote-Identity-Proofing-Attacks-Countermeasures-1.pdf

ENISA. (2024). Remote ID proofing good practices. In E. Nikolouzou & R. Naydenov (Eds.), *ENISA report.* Publications Office of the European Union. ISBN: 978-92-9204-661-3. https://www.enisa.europa.eu/publications/remote-id-proofing-good-practices

Europol. (2022). Facing reality? Law enforcement and the challenge of deepfakes. *Publications Office of the European Union.* ISBN: 978-92-95236-23-3. https://op.europa.eu/s/z12G

Garrido, P., Valgaerts, L., Rehmsen, O., Thorm, T., Perez, P., & Theobalt, C. (2014). Automatic face reenactment. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2014).* IEEE. ISBN: 978-1-4799-5118-5. http://dx.doi.org/10.1109/CVPR.2014.537.

Grand View Research. (2020). Jelszó nélküli hitelesítési piacméret, részesedés és trendelemzési jelentés komponensenként, terméktípusonként (arcfelismerés, ujjlenyomat, írisz), hitelesítési típus szerint, hordozhatóság szerint, végfelhasználó szerint, régió szerint és szegmens-előrejelzések szerint. *Market Analysis Report 2022–2030.* GVR-4-68039-996-6. https://www.grandviewresearch.com/industry-analysis/passwordless-authentication-market-report

Hernandez-Ortega, J., Fierrez, J., Morales, A., & Galbally, J. (2019). Introduction to face presentation attack detection. In S. Marcel, J. Fierrez, & N. Evans (Eds.), *Handbook of biometric anti-spoofing* (pp. 203–230). Springer. https://doi.org/10.1007/978-3-319-92627-8_9

Kramer, R. S. S., Mireku, M. O., Flack, T. R., & Kay, L. R. (2019). Face morphing attacks: Investigating detection with humans and computers. *Cognitive Research: Principles and Implications, 4*(28). https://doi.org/10.1186/s41235-019-0181-4

Ngan, M., Grother, P., & Hanaoka, K. (2022). NIST IR 8331 DRAFT SUPPLEMENT: Ongoing Face Recognition Vendor Test (FRVT) Part 6B: Face recognition accuracy with face masks using post-COVID-19 algorithms. *National Institute of Standards and Technology (NIST)*. https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt-ongoing

Ngan, M., Grother, P., & Hom, A. (2023). NIST Internal Report NIST IR 8491: Face analysis technology evaluation (FATE), Part 10: Performance of passive, software-based presentation attack detection (PAD) algorithms. *National Institute of Standards and Technology (NIST).* https://doi.org/10.6028/NIST.IR.8491

Ngan, M., Grother, P., Hanaoka, K., & Kuo, J. (2024). NIST IR 8292 DRAFT SUPPLEMENT: Face analysis technology evaluation (FATE), Part 4: MORPH - Performance of automated face morph detection. *National Institute of Standards and Technology (NIST)*. https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt-ongoing

Qin, L., Peng, F., Long, M., Ramachandra, R., & Busch, B. (2021). Vulnerabilities of unattended face verification systems to facial components-based presentation attacks: An empirical study. *ACM Transactions on Privacy and Security, 25*(1), Article 4, pp. 1–28. https://doi.org/10.1145/3491199

Rathgeb, C., Tolosana, R., Vera-Rodriguez, R., & Busch, C. (Eds.). (2024). *Handbook of digital face manipulation and detection: From DeepFakes to morphing attack.* Springer. https://doi.org/10.1007/978-3-030-87664-7

Venkatesh, S., Ramachandra, R., Raja, K., & Busch, C. (2021). Face morphing attack generation and detection: A comprehensive survey. *IEEE Transactions on Technology and Society, 2*(3), pp. 128–145. https://ieeexplore.ieee.org/document/9380153

Yan, Z., Zhang, Y., Yuan, X., Lyu, S., & Wu, B. (2023). DeepfakeBench: A Comprehensive Benchmark of Deepfake Detection. *Proceedings of the 37th Conference on Neural Information Processing Systems (NeurIPS 2023), Track on Datasets and Benchmarks*. https://arxiv.org/pdf/2307.01426

Yu, Z., Qin, Y., Li, X., Zhao, C., Lei, Z., & Zhao, G. (2023). Deep learning for face anti-spoofing: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 45*(5), pp. 5609–5631. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9925105

Zhao, Z., Wang, P., & Lu, W. (2021). Multi-layer fusion neural network for deepfake detection. *International Journal of Digital Crime and Forensics, 13*(4), pp. 26–39. http://dx.doi.org/10.4018/IJDCF.20210701.oa3

Alkalmazott jogszabályok

Az EU 2024/1689 rendelete (AI Act, mesterséges intelligencia rendelet)

Az EU 2024/1183 rendelete (eIDAS 2.0 rendelet)

A cikk APA szabály szerinti hivatkozása

Hazai L. (2025). Az automatikus biometrikus azonosság-ellenőrző rendszerek védelmének kérdései – Prezentációs támadások, megtévesztés. *Belügyi Szemle*, *73*(3), 477–494. https://doi.org/10.38146/BSZ-AJIA.2025.v73.i3.pp477-494

Nyilatkozatok

Összeférhetetlenség

A szerző nem jelentett összeférhetetlenséget.

Finanszírozás

A szerző nem kapott pénzügyi támogatást a kutatáshoz, a szerzőséghez és/vagy a cikk publikálásához.

Etikai nyilatkozat

Jelen cikkhez nem kapcsolódik adatkészlet.

Levelező szerző

A cikk levelező szerzője Hazai Lászlóné, aki a hazai.laszlone@nbsz.gov.hu e-mail címen érhető el.