OUCH! 2022. március

Sajátítsunk el egy új túlélési készséget: ismerjük fel a deepfake-eket!

Mik azok a deepfake-ek?

A „deepfake” szóösszetétel a „deep learning” (mély gépi tanulás) és a „fake” (hamis) kifejezésekből áll. A deepfake-ek gyakorlatilag hamisított képek, videók vagy hanganyagok. Esetenként az ezekben szereplő emberek – bár valós személyeknek tűnnek – igazából nem is léteznek, csupán számítógépes animációval hozták létre őket. Az is előfordulhat, hogy létező személyek megjelenését, vagy hangját manipulálják, és a felvételen olyan dolgot tesznek, vagy mondanak, ami a valóságban nem történt meg. Például egy deepfake videóban annak készítői akár egy jól ismert híresség vagy politikus is megjeleníthetnek, olyan szöveget adva a szájába, amit soha sem állított. Az ilyen életszerű deepfake anyagokkal a támadók egy alternatív valóságot termethetnek, amelyben nem hihetünk sem a szemünknek, sem a fülünknek.

Egyes deepfake-eknek ugyanakkor legitim funkciójuk is lehet, például mozifilmek esetében egy korábban elhunyt színész emlékének feleleventésével. Azonban a kibertámadók elkezdtek visszaélni a depfake technológia adta lehetőségekkel. A támadók sok esetben azért alkalmazzák ezt a technikát, hogy megtévesszenek bennünket, ellophassák pénzünket, zaklassanak, befolyásolják politikai nézetünket vagy épp azért, hogy álhíreket kreáljanak. Előfordult már olyan is, hogy komplett kamu vállalatot készítettek, amelyet deepfake karakterekkel töltöttek fel. Az ilyen támadások fényében érdemes kétszer is átgondolni, hogy higgyünk-e egy-egy hírnek vagy közösségi média posztnak.

Az FBI nemrég arra figyelmezetett, hogy a jövőben a deepfake-ek még erőteljesebb hatással bírnak majd, a technológia fejlődése miatt. Tanuljuk meg felismerni a deepfake-eket, hogy képesek legyünk védekezni azok káros hatásai ellen! A deepfake minden formája – legyen az egy állókép, videó, vagy csupán hang – olyan sajátos hibákkal rendelkezik, amelyek lebuktathatják.

Állóképek

A leggyakoribb deepfake tartalom, amivel találkozhatunk, a kamu közösségi média profilkép. Az alábbi kép egy konkrét példa deepfake-re a thispersondoesnotexist.com című weboldalról. A kép alatt öt olyan nyom található, amelyek arra utalhatnak, hogy ez egy deepfake lehet. Láthatjuk, hogy ezeknek a nyomoknak a felismerése eseteknként nem is olyan könnyű:

1. Háttér: A háttér sok esetben homályos vagy torz, és a kép fényessége sem egyenletes, például az árnyékok nem természetesek.
2. Üvegfelületek: Vizsgáljuk meg figyelmesen a szemüvegkeret és a homlok közötti átmenetet. A deepfake-ek egyes részei esetenként nem illeszkednek egymáshoz, méret- vagy formabeli eltérések észlelhetők.
3. Szemek: A deepfake fotókra jelenleg jellemző az „üveges tekintet”.
4. Ékszerek: A fülbevalók amorfak lehetnek vagy a rögzítésük furcsának tűnhet. A nyakláncok, mintha a bőrbe lennének ágyazva.
5. Nyak és vállak: A vállak alaktalanok vagy egyszerűen nem passzolnak. A nyak különbözőnek hat a két oldalon.

Videó

Az MIT (Massachusetts Institute of Technology) kutatói elkészítettek egy kérdéssort, amely segítségünkre lehet abban, hogy kiderítsük egy videó valódi-e, kiemelve, hogy a deepfake-ek gyakran nem keltenek természetes hatást az ábrázolt helyszín, illetve a fényesség tekintetében.

1. Arc és homlok: Túl simának vagy túl ráncosnak tűnik a bőr? Ugyanolyan idősnek tűnik a bőr, mint a haj, vagy a szemek?
2. Szem és a szemöldök: Az árnyékok ott jelennek meg, ahol számítunk rájuk?
3. Üvegfelületek: Van tükröződés? Vagy éppen túl sok a tükröződés? Változik-e a tükröződés szöge, amikor a személy mozog?
4. Arcszőrzet: Valódinak tűnik? A deepfake-ekben sok esetben hozzáadnak vagy elvehetnek bajuszt, pajeszt vagy szakállt.
5. Arcon lévő anyajegyek: Valódinak tűnnek?
6. Pislogás: A szereplő eleget pislog vagy éppen túl sokat?
7. Az ajkak mérete és színe: A méret és a szín megegyezik a személy arcának többi részével?

Hang

A kutatók szerint az olyan technológiák, mint a spektrogramok, képesek jelezni, ha a hangfelvételek hamisak. Azonban a legtöbbünknek nincs kéznél hangelemző, amikor egy támadó telefonál. Vegyük észre, ha túl monoton az előadást, a furcsa hangmagasságot vagy érzelmeket, valamint a háttérzaj hiányát. A hanghamisításokat nehéz felismerni. Ha furcsa hívást kapunk egy szervezettől, ellenőrizhetjük, hogy a hívás valódi-e, ha először letesszük a kagylót, majd visszahívjuk a szervezetet. Ügyeljünk arra, hogy megbízható telefonszámot használjunk, például olyat, amely szerepel a telefonos névjegyzékünkben, a szervezettől kapott számlán vagy a szervezet hivatalos webhelyén.

Következtetés

Legyünk tudatában annak, hogy a támadók aktívan alkalmazzák a deepfake technológiát! Hamis fiókokat hozhatnak létre a közösségi médiában, hogy kapcsolatba lépjenek velük, vagy hamis videókat készíthetnek a közvélemény befolyásolása érdekében. Néhányan a sötét weben bérbe is adják a szolgáltatásaikat, így más támadók is megtehetik ugyanezt. Nem várjuk el, hogy deepfake szakértővé váljon, de ha a hamisítványok azonosításához szükséges tudás alapjaival tisztában lesz, sokkal könnyebben lesz képes védekezni. Ha azt gyanítja, hogy deepfake-kel találkozott, jelentse azt a tartalmat tároló webhelynek vagy forrásnak.

A szerzőről

Kerry Tomlinson (@KerryTNews) az Ampere News kiberbiztonsági riportere, aki egyben SANS Security Awareness tanúsítvánnyal rendelkező szakember. Küldetésének tekinti, hogy a digitális világ történéseit lefordítsa a különböző tudással rendelkezők számára.

Források

Pszichológiai manipulációs támadások: https://www.sans.org/newsletters/ouch/social-engineering-attacks/
Ki tudja szúrni, hogy mi a hamis? Ampere News: https://www.amperesec.com/news/can-you-spot-the-fake
Az MIT deepfake felismerésére szolgáló tesztje: https://detectfakes.media.mit.edu/
Szúrjuk ki a deepfake-et: https://www.spotdeepfakes.org/en-US

A fordítást készítette: Nemzetbiztonsági Szakszolgálat Nemzeti Kibervédelmi Intézet (NBSZ NKI)