OUCH! 2025. március

Az Ön Havi Biztonsági Tudatosságról szóló hírlevele

Óvakodjon a deepfake-től: a megtévesztések új korszaka

Váratlanul ért: Steve története

Steve az asztalánál ült, amikor a főnöke, Bela, zaklatottan felhívta. A videóhívásban a hölgy idegesnek tűnt, a hangja sietős volt. “Most azonnal küldd el a bizalmas ügyféljelentést az új email címemre!” kérte. Látva az ismerős arcot és hallva a jellegzetes hangját, Steve nem hezitált, azonnal elküldte a bizalmas jelentést az új email címre.

Pár órával később Bela besétált a férfi irodájába és a jelentés után érdeklődött. Steve összezavarodva megemlítette a korábbi videóhívást. Bela megdöbbent – hiszen ő nem is hívta Steve-et. Az ember, akit a férfi a videón látott, nem Bela volt. Egy *deepfake* volt, amelyet egy kiberbűnöző hozott létre, hogy átverje őt.

Steve alig hitte el, milyen valóságosnak tűnt a hamis hívás. Az arc és a hang is a főnökének tökéletes mása volt. Ő is áldozatul esett annak az egyre népszerűbb kiberfenyegetésnek, amikor a bűnözők mesterséges intelligenciát (MI) használnak minél valósághűbb másolatok létrhozására.

Mi az a deepfake?

Az MI képes olyan képeket, hanganyagokat és videókat létrehozni, amelyek valódinak tűnnek. Ezen képességek számos legitim felhasználásra alkalmasak. Példának okáért a marketing cégek arra használják a technológiát, hogy képeket hozzanak létre a reklámkampányaikhoz, a filmipar ezzel fiatalítja a színészeket, a tanárok pedig ennek segítségével hozhatnak létre dinamikus videókat a tananyagaikhoz.

A deepfake az, amikor az MI segítségével készítenek hamis képeket, hanganyagokat, videókat azzal a céllal, hogy másokat átverjenek. A “deepfake” név a “deep learning”, magyarul “mélytanulás” (egy MI típus) és a “fake” (azaz “hamis”) szavak kombinációjából ered.

A legveszélyesebb deepfake-ek gyakran azok, amikor a kiberbűnözők hamis képeket, hangfelvételeket vagy videókat készítenek ismerős emberekről, és olyan helyzetekben ábrázolják őket, amelyek soha nem történtek meg. Például olyan hamis képeket alkothatnak híres celebekről vagy politikusokról, amelyekben valamilyen bűnt követnek el, és hamis hírekként terjeszthetik. Vagy lemásolhatják valaki hangját, és felhasználhatják egy hívásban, hogy megtévesszék az áldozat családját vagy kollégáit. A deepfake-ek azért különösen veszélyesek, mert a támadók könnyedén leképezhetnek bárkit, a létehozott anyagokkal bármit megtehetnek, és úgy tűntethetik fel, mintha valósak lennének.

A deepfake három típusa

1. Deepfake képek

Ezek vagy mesterséges intelligencia által létrehozott hamis emberekről készült képek, vagy valódi emberekről készült fotók, amelyek olyan helyzeteket ábrázolnak, amelyek soha nem történtek meg. Ezek a hamis képek gyorsan terjedhetnek és gyakran hitelrontásra vagy érzelmek manipulálására használják őket. A deepfake képek száma egyre inkább növekszik a közösségi médiában, és emberek vagy akár kormányok is igyekeznek hamis történeteket illetve narratívákat (azaz fake news-t) terjeszteni egy bizonyos cél elérése érdekében.

2. Deepfake audiók (hang klónozás)

Ezek olyan hamis hangfelvételek vagy telefonhívások, amelyek valaki hangjának klónozásával készülnek. A támadók megszerezhetik valakinek a hangfelvételeit például podcastekből vagy YouTube-ról, majd ezeket felhasználva lemásolhatják a hangját. A leképzést követően a támadók felhívhatnak bárkit akit szeretnének, miközben a lemásolt illetőnek adják ki magukat. Például egy támadó vezetőnek adhatja ki magát, és felhívhat egy alkalmazottat, hogy bizalmas információkat szerezzen meg, vagy egy családtag hangját lemásolva vészhelyzetre hivatkozva pénzt kérhet.

3. Deepfake videók

Ezek hamis videók, amelyekben személyek hangját és testmozdulatait utánozzák le vagy manipulálják. Ezek az audiovizuális anyagok előre felvett vagy élő videók is lehetnek, mint például akár egy online konferenciahívás is. A kibertámadók például készíthetnek egy olyan deepfake videót, amelyben egy vezérigazgató hamis bejelentést tesz a cégével kapcsolatban, vagy egy politikus olyan kijelentést mond, ami valójában soha nem történt meg.

Hogyan ismerhetjük fel a deepfake tartalmakat: figyeljünk a kontextusra

Ne próbáljuk meg felismerni a deepfake-eket pusztán technikai hibákat keresve! Mind az MI, mind pedig a kibertámadók rendkívül kifinomulttá váltak. Ehelyett figyeljünk a kontextusra. A képnek, hanganyagnak vagy videónak van bármi értelme?

1. Bízzunk az ösztöneinkben: Furcsának érződik valami? A kérés sürgős vagy váratlan? Az illető a külseje és a hangja alapján normálisnak tűnik, mégis van valami szokatlan a viselkedésében? Valaki olyan bizalmas információt vagy személyes adatot kér, amelyhez nem szabadna hozzáférnie? Ha valami gyanúsnak tűnik, hallgassunk az ösztöneinkre, és ellenőrizzük le alaposan, mielőtt teljesítenénk a kérést!

2. Óvakodjunk az érzelmi zsarolástól: A támadók gyakran siettetik vagy megijesztik áldozataikat, a gyors cselekvés érdekében. Ha egy üzenet vagy hívás pánikot kelt bennünk, vegyünk egy mély levegőt és ellenőrizzük! Minél erősebb az érzelmi hatás, például ha sürgető a helyzet vagy félelmet keltő, annál nagyobb az esélye, hogy egy támadásról van szó.

3. Ellenőrizzük más módszerekkel: Ha attól tartunk, hogy a kapcsolatfelvevő személy deepfake lehet, próbáljuk meg elérni őt egy másik kommunikációs csatornán! Ha mondjuk attól tartunk, hogy egy videóhívás vagy üzenet hamis, vegyük fel a kapcsolatot a másik féllel telefonon vagy emailben! Ha olyan hívást kapunk, amely azonnali cselekvésre buzdít, tegyük le a telefont és hívjuk vissza egy megbízható telefonszámon!

4. Találjunk ki egy "kódszót" vagy mondatot: Állapodjunk meg egy közös jelszóban vagy kifejezésben, amelyet csak a csoport vagy a család tagjai ismernek, és amelyet sürgős üzenetek hitelesítésére használhatunk!

Vendégszerkesztő

Dhruti Mehta információbiztonsági szakértő a Physicians Health Plan of Norhtern Indiana-nal, illetve az észak indianai WiCyS elnöke. Szenvedélye a sokszínű munkaerő kiépítése, valamint az oktatási és készségbeli hiányosságok áthidalása a kiberbiztonság területén. https://www.linkedin.com/in/dhrutimehtacyber/

Források

Érzelmi triggerek – Így csapnak be minket a kibertámadók: https://www.sans.org/newsletters/ouch/emotional-triggers-how-cyber-attackers-trick-you
Hang klónozás: https://www.sans.org/newsletters/ouch/phantom-voices-defend-against-voice-cloning-attacks/

A Közösség számára fordította: Nemzetbiztonsági Szakszolgálat Nemzeti Kibervédelmi Intézet (NBSZ NKI)