OUCH! 2025. gada marts

Jūsu ikmēneša informatīvais biļetens drošības izpratnes veicināšanai

Uzmanieties no dziļviltojumiem: Jauns maldināšanas laikmets

Apkrāpts negaidīti: Stīva stāsts

Stīvs sēdēja pie rakstāmgalda, kad saņēma izmisīgu video zvanu no savas vadītājas Bellas. Video zvanā viņa izskatījās uztraukusies, viņas balss bija steidzīga. "Man vajag, lai tu nekavējoties nosūtītu konfidenciālu klienta ziņojumu uz šo jauno e-pastu!" viņa uzstāja. Redzot viņas pazīstamo seju un dzirdot viņai raksturīgo balsi, viņš nevilcinājās un nosūtīja konfidenciālo ziņojumu uz jauno e-pasta adresi.

Dažas stundas vēlāk viņa birojā ienāca Bella un jautāja par ziņojumu. Apjukušais Stīvs pieminēja video zvanu. Bellas seja kļuva līķa bāla – viņa nebija zvanījusi Stīvam. Persona, kuru viņš redzēja video zvanā, nebija Bella. Tas bija *dziļviltojums*, ko radījis kibernoziedznieks, lai viņu apmānītu.

Stīvs nespēja noticēt – viltus zvans šķita tik reāls. Seja, balss – viss perfekti atbilda viņa vadītājai. Viņš bija kļuvis par upuri pieaugošajam kiberapdraudējumam, kad noziedznieki izmanto mākslīgo intelektu (MI), lai radītu ļoti pārliecinošus viltojumus.

Kas ir dziļviltojums?

Mākslīgais intelekts var radīt attēlus, audio vai video, kas izskatās atbilstoši realitātei. Šīm iespējām ir daudz likumīgu pielietojumu. Piemēram, mārketinga uzņēmumi izmanto šo tehnoloģiju, lai radītu attēlus reklāmas kampaņām, kino industrija to izmanto, lai atdarinātu noteiktus aktierus, un skolotāji to izmanto, lai radītu dinamiskas video nodarbības saviem skolēniem.

Dziļviltojums ir, kad mākslīgais intelekts tiek izmantots, lai radītu viltotus attēlus, audio vai videoklipus ar mērķi maldināt citus. Nosaukums "dziļviltojums" radies no angļu valodas vārda "deepfake", divu vārdu salikuma: "deep learning" (dziļmācīšanās – mākslīgā intelekta veids) un "fake" (viltojums) kombinācijas

Bieži vien viskaitīgākie dziļviltojumi ir, kad kibernoziedznieki rada viltotus attēlus, audio vai video ar jums pazīstamiem cilvēkiem, panākot, ka viņi dara lietas, ko patiesībā nekad nav darījuši. Piemēram, kiberuzbrucēji var radīt viltotus attēlus, kuros redzamas slavenības vai politiķi, kas izdarījuši noziegumu, un izplatīt tos kā viltus ziņas. Vai arī viņi var klonēt kāda cilvēka balsi un izmantot to zvanā, lai maldinātu upura ģimeni vai kolēģus. Sevišķi bīstami ir tas, cik viegli kibernoziedznieki var atdarināt jebkuru personu, panākot, ka tā dara jebko, un likt tai izskatīties reālai.

Trīs dziļoviltojumu veidi

1. Attēla dziļviltojumi

Tie ir vai nu mākslīgā intelekta radīti viltus cilvēku attēli, vai arī reālu cilvēku attēli, kuros redzams, ka viņi dara kaut ko tādu, ko viņi nekad nav darījuši. Šie viltus attēli var ātri izplatīties, un tos bieži izmanto, lai kaitētu reputācijai vai manipulētu ar emocijām. Sociālajos medijos arvien biežāk tiek izmantoti viltus attēli, un cilvēki vai pat valdība cenšas virzīt viltotus stāstus vai naratīvus (tā dēvētās viltus ziņas), lai panāktu vēlamo mērķi.

2. Audio dziļviltojumi (balss klonēšana)

Tie ir viltus ieraksti vai tālruņa zvani, kuros izmantota kāda cilvēka klonēta balss. Uzbrucēji var iegūt cilvēku balsu ierakstus no WhatsApp audio ziņām, ja tiek uzlauzts konts, YouTube vai citiem sociālajiem medijiem, kuros ir brīvi pieejami video un audi ieraksti ar personu. Pēc tam šie ieraksti tiek izmantoti, lai atdarinātu, piemēram, personas balsi. Pēc ieraksta iegūšanas un apstrādes (replikācijas) kiberuzbrucēji var piezvanīt jebkurai personai, izliekoties par šo personu. Piemēram, kāds var izlikties par vadītāju un piezvanīt darbiniekam, lai pieprasītu sensitīvus datus, vai arī kāds var atveidot mīļotā cilvēka balsi, zvanot ārkārtas situācijā un lūdzot naudu.

3. Video dziļviltojumi

Tie ir viltus videoklipi, kuros cilvēku balss un darbības tiek manipulētas vai atveidotas. Video dziļviltojums var būt iepriekš ierakstīts video vai tiešraides video, piemēram, tiešsaistes konferences zvans. Piemēram, kiberuzbrucēji var izveidot video veida dziļviltojumu, kurā uzņēmuma vadītājs sniedz viltus paziņojumu par savu uzņēmumu vai politiķis šķietami saka kaut ko tādu, ko viņš nekad nav teicis.

Kā atklāt dziļviltojumus: koncentrējieties uz kontekstu

Nemēģiniet atklāt dziļviltojumus, meklējot tehniskas kļūdas. Gan mākslīgais intelekts, gan kiberuzbrucēji, kas to izmanto, ir kļuvuši ļoti prasmīgi. Tā vietā pievērsiet uzmanību kontekstam. Vai attēlam, audio vai video ir jēgpilns?

1. Uzticieties saviem instinktiem: Vai komunikācijā kaut kas šķiet nepareizi? Vai lūgums ir steidzams vai negaidīts? Vai persona uzvedas dīvaini, pat ja tā izskatās un izklausās normāli? Vai kāds pieprasa konfidenciālu informāciju vai personas datus, kuriem viņam nevajadzētu piekļūt? Ja kaut kas nešķiet pareizi, paļaujieties uz savu intuīciju un vēlreiz pārbaudiet, pirms izpildāt šo lūgumu.

2. Uzmanieties no emocionālās manipulācijas: Kiberuzbrucēji bieži vien rada steidzamību vai bailes, lai piespiestu jūs rīkoties ātri. Ja ziņa vai zvans rada paniku, paņemiet pauzi un pārbaudiet informāciju. Jo spēcīgāka ir emocionālā ietekme, piemēram, radīta spēcīga steidzamības vai baiļu sajūta, jo lielāka ir varbūtība, ka tas ir potenciāls uzbrukums.

3. Pārbaudiet, izmantojot citu metodi: Ja jums ir bažas, ka persona, kas ar jums sazinās, varētu būt viltus persona, sazinieties ar šo personu, izmantojot citu metodi. Piemēram, video zvanu vai ziņojumu gadījumā, par kuriem jums ir bažas, ka tie varētu būt maldinoši, sazinieties ar personu pa tālruni vai e-pastu. Ja saņemat balss zvanu ar aicinājumu steidzami rīkoties, nolieciet klausuli un zvaniet atpakaļ, izmantojot uzticamu numuru.

4. Izveidojiet koda vārdu vai frāzi: vienojaties par kopīgu koda vārdu vai frāzi, kas zināma tikai konkrētajā grupā vai jūsu ģimenē un ko varat izmantot, lai apliecinātu savu identitāti situācijās, kad nepieciešama ātra rīcība.

Viesredaktore

Druti Mehta (Dhruti Mehta) ir informācijas drošības analītiķe Ziemeļindiānas veselības aprūpes plānā (Physicians Health Plan of Northern Indiana) un WiCyS Northern Indiana prezidente. Viņa aizrautīgi strādā daudzveidīga kiberdrošības darbaspēka veidošanā un izglītības un prasmju trūkuma mazināšanā šajā jomā. https://www.linkedin.com/in/dhrutimehtacyber/

Resursi

Emocionālie ierosinātāji – kā kiberuzbrucēji piemāna cilvēkus: https://www.sans.org/newsletters/ouch/emotional-triggers-how-cyber-attackers-trick-you/
Balss klonēšana: https://www.sans.org/newsletters/ouch/phantom-voices-defend-against-voice-cloning-attacks/

CERT.LV