As deepfake são uma real ameaça para todos nós

dinheirovivo.pt/opiniao/as-deepfake-sao-uma-real-ameaca-para-todos-nos-17484125.html

11 de dezembro de 2023

A rápida evolução da era digital, inaugurou uma nova fase de preocupação e escrutínio, onde os *deepfakes* aparecem numa posição central. Os *deepfakes*, alimentados por algoritmos sofisticados de inteligência artificial, permitem a criação de meios sintéticos hiper-realistas, incluindo vídeos manipulados e gravações áudio. Embora estas tecnologias sejam promissoras em vários domínios, desde o entretenimento à produção virtual, as suas aplicações maliciosas suscitam preocupações de segurança significativas.

Os *deepfakes* podem ser utilizados para espalhar desinformação e notícias falsas. Ao criar vídeos realistas ou gravações áudio de figuras públicas a dizer ou a fazer coisas que nunca fizeram, há margem para manipular a opinião pública e criar narrativas falsas.

Mas há mais.

A natureza realista do conteúdo *deepfake* torna mais difícil para o público em geral, o discernimento entre a informação autêntica e a manipulada.

Os *deepfakes* podem também ser utilizados como uma ferramenta na guerra cibernética. Imagens realistas de indivíduos de alto nível ou funcionários governamentais têm o potencial de criar tensões diplomáticas e confusão geopolítica.

Agravam ainda as tensões sociais através da criação de conteúdos fabricados que alimentam conflitos ou retratam os acontecimentos de forma enganadora. Esta situação pode ter consequências no mundo real e provocar agitação social. Assim, os *deepfakes* representam também uma ameaça inegável à segurança nacional, uma vez que podem ser explorados para manipular cenários políticos. Por exemplo, através da criação de vídeos falsos de líderes políticos a proferirem declarações controversas ou a adotar comportamentos inadequados, pode influenciar a perceção do público acabando por também influenciar e até manipular, resultados eleitorais.

Da mesma forma, os cibercriminosos podem utilizar *deepfakes* para criar conteúdos convincentes, levando os seus alvos a acreditar que estão a interagir com uma pessoa de confiança, o que conduz a fraudes financeiras ou ao acesso não autorizado a informações sensíveis.

As empresas não estão imunes a esta ameaça. Os líderes e executivos das empresas podem tornar-se alvos fáceis, com os *deepfakes* a serem utilizados para manipular os mercados financeiros, perturbar as operações comerciais ou espalhar desinformação que prejudique a reputação das empresas.

Adicionalmente, a autenticidade dos *deepfakes* pode ser aproveitada em ataques de engenharia social. Tanto os *deepfakes* como esta, dependem da manipulação da confiança. Permitem criar o conteúdo que a engenharia social explora para enganar os alvos e os leva a tomar determinadas ações. Assim, a combinação da tecnologia *deepfake* com as técnicas de engenharia social tem o potencial de aumentar a eficácia do engano, apresentando um conteúdo altamente realista e genuíno. Quando combinados com táticas de engenharia social, os *deepfakes* podem iludir e aumentar o engano, fornecendo provas visuais ou auditivas para apoiar a narrativa manipuladora.

A falsificação de identidade é uma tática comum de engenharia social e os *deepfakes* podem ser utilizados para criar imitações realistas, sobrepondo o rosto de alguém ao corpo de outra pessoa num vídeo. Este elemento visual aumenta a eficácia das tentativas de falsificação de identidade.

Os ataques envolvem frequentemente vários canais de comunicação, como emails, chamadas telefónicas e mensagens, onde os *deepfakes* podem ser facilmente integrados, possibilitando a apresentação de uma mensagem consistente e convincente em diferentes meios.

Todavia, é de realçar que estas falsificações podem ir muito mais longe do que a simples troca de rostos num vídeo. O áudio *deepfake* também está a crescer e tem causado algumas falsificações muito convincentes. Através dos mesmos algoritmos de *deep learning*, a voz de alguém pode ser replicada de forma muito realista.

Por exemplo, um vídeo ou mensagem de áudio *deepfake* de uma pessoa de confiança pode ser utilizado num ataque de engenharia social para aumentar a probabilidade de sucesso. Tal como, de facto, já aconteceu: durante um ciberataque, os atores, fazendo-se passar pelo CEO de uma empresa, enviaram um *email* com um ficheiro de áudio no qual pediam uma transação de milhões de euros, transação essa que foi efetuada. O *software* foi capaz de imitar a voz, e não só a voz: a tonalidade, a pontuação e o sotaque. Este, entre outros, é um exemplo de como a engenharia social se tem vindo a tornar altamente complexa e de como os *deepfakes* ajudam ao seu sucesso.

O sucesso da engenharia social depende muitas vezes da credibilidade da história ou da personalidade do atacante. Manipular a perceção é um aspeto fundamental e os *deepfakes* aumentam a sua credibilidade fornecendo provas visuais ou auditivas que apoiam a narrativa da engenharia social.

Finalizo com algo que gosto de vincar: segurança baseia-se na confiança. Os *deepfakes* proporcionam confiança onde ela não deveria existir. A simples realidade é que a tecnologia de produção de *deepfake* está atualmente a melhorar mais rapidamente do que a tecnologia de deteção de *deepfake*.