UBND TỈNH LONG AN
SỞ GIÁO DỤC VÀ ĐÀO TẠO

Số: 3750 /SGDĐT-HCQT
V/v thông tin 1 số thủ đoạn sử dụng trí tuệ nhân tạo nhằm mục đích phạm tội

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

Long An, ngày 27 tháng 9 năm 2024

Kính gửi:
- Trưởng phòng GD&ĐT các huyện, thị xã, thành phố;
- Thủ trưởng các đơn vị trực thuộc Sở.

Căn cứ Công văn số 2870/CAT-PA05 ngày 04/9/2024 của Công an tỉnh về việc thông tin một số thủ đoạn sử dụng trí tuệ nhân tạo nhằm mục đích phạm tội.

Sở GD&ĐT thông tin đến các đơn vị về một số thủ đoạn sử dụng trí tuệ nhân tạo nhằm mục đích phạm tội theo Công văn 2870/CAT-PA05 ngày 04/9/2024 như sau:

Trong thời gian qua, tình hình tội phạm sử dụng AI diễn ra hết sức phức tạp với nhiều thủ đoạn tinh vi. So với loại hình phạm tội truyền thống, tội phạm sử dụng AI hoạt động chủ yếu trên không gian mạng, thường xuyên thay đổi phương thức, thủ đoạn, gây khó khăn cho cơ quan chức năng trong việc phát hiện và xử lý. Một số thủ đoạn sử dụng trí tuệ nhân tạo nhằm mục đích phạm tội phổ biến hiện nay bao gồm:

1. Sử dụng công nghệ AI để tạo ra các hình ảnh, video không có thật:

Deepfake là công nghệ được các đối tượng chủ yếu sử dụng nhằm mục đích tạo ra các hình ảnh, video không có thật trong thời gian qua. Một số hành vi phạm tội sử dụng công nghệ Deepfake, cụ thể:

- Truyền bá thông tin sai lệch, tin giả nhằm ảnh hưởng đến dư luận, gây bất ổn xã hội: Tội phạm sử dụng Deepfake tạo ra những tài liệu không có thật với mục đích tác động đến dư luận xã hội.

- Làm tổn hại đến danh tiếng, uy tín hoặc cưỡng đoạt tài sản bằng cách mạo danh, bôi nhọ hình ảnh của cá nhân, tổ chức: Nạn nhân thường là người nổi tiếng, có sức ảnh hưởng trên các nền tảng mạng xã hội như facebook, tiktok, instagram, đăng tải nhiều hình ảnh cá nhân, riêng tư trên không gian mạng. Các đối tượng sử dụng hình ảnh trên mạng xã hội, thông qua Deepfake để tạo ra các hình ảnh, video có mặt của nạn nhân, chứa nội dung "nhạy cảm", không có thật hoặc mạo danh nạn nhân để thực hiện các hành vi không chuẩn mực, gây ảnh hưởng đến danh dự của nạn nhân, từ đó cưỡng đoạt tài sản của nạn nhân.

- Lừa đảo chiếm đoạt tài sản: Các đối tượng sử dụng Deepfake giả dạng cá nhân, tổ chức để đe dọa, yêu cầu nạn nhân chuyển tiền hoặc cung cấp mã OTP để phá vỡ lớp bảo mật, sau đó chiếm đoạt tiền từ tài khoản của nạn nhân. Một số thủ đoạn bao gồm: Giả danh cơ quan chức năng để đe dọa nạn nhân: Các đối tượng giả mạo là cán bộ cơ quan nhà nước (Công an, Viện kiểm sát, Tòa án...) gọi điện trực tuyến với nạn nhân, thông báo ghi nhận hình ảnh nạn nhân vi phạm pháp luật; lợi dụng tâm lý sợ hãi, không muốn bị mang tiếng xấu của nạn nhân để lừa đảo, yêu cầu nạn nhân chuyển tiền để giải quyết vụ việc. Giả mạo người thân của nạn nhân để lừa đảo, tội phạm sử dụng Deepfake để tạo ra những video, hình ảnh giả là người thân của nạn nhân để tạo niềm tin của nhạn nhân, từ đó yêu cầu nạn nhân "giúp đỡ", chuyển tiền vào số tài khoản "lạ" và chiếm đoạt tài sản.

- Vượt qua xác minh danh tính: đối tượng sử dụng Deepfake giả mạo danh tính khuôn mặt nạn nhân để vượt qua xác minh sinh trắc học mở khóa bằng khuôn mặt trên các thiết bị di động, từ đó chiếm quyền sử dụng thiết bị di động của nạn nhân.

- Tấn công mạng: Đã xảy ra một số vụ tấn công mạng sử dụng Deepfake, trong đó email là phương thức được sử dụng thường xuyên, nhằm tạo ra nội dung giả mạo và gửi nội dung đó đến "trình phân biệt đối xử" của hệ thống, sau đó so sánh nội dung giả mạo với nội dung thật để xác minh sự khác biệt giữa hai nội dung này và loại bỏ những khác biệt, đánh lừa "trình phân biệt đối xử" một lần nữa thông qua nội dung giả mạo đã được cải tiến. Chu kỳ này tiếp tục cho đến khi một tệp giả mạo gần như hoàn hảo được tạo ra.

2. Sử dụng AI trong phương tiện không người lái nhằm mục đích phạm tội:

- Buôn lậu ma túy, vũ khí, hàng cấm: Tội phạm sử dụng phương tiện không người lái để vận chuyển hàng buôn lậu đến địa điểm đã được định vị sẵn.

- Buôn người, đưa người vượt biên trái phép: tương tự với hoạt động buôn lậu ma túy, vũ khí, hàng cấm, tội phạm sử dụng phương tiện không người lái làm công cụ vận chuyển người để tránh sự điều tra của cơ quan chức năng trong trường hợp bị bắt giữ.

- Tấn công khủng bố: Tội phạm sử dụng phương tiện không người lái được lập trình để tấn công các khu vực đông đúc hoặc mục tiêu đã được chỉ định sẵn, thậm chí có thể kích hoạt các thiết bị gây nổ nhằm gây thiệt hại lớn về người và tài sản.

- Hoạt động gián điệp: phương tiện không người lái tích hợp AI được tội phạm sử dụng trong hoạt động lấy cắp thông tin qua các hình ảnh mà nó chụp được, hoặc được sử dụng để tấn công hệ thống máy tính.

3. Sử dụng AI tạo ra các mã độc, phần mềm độc hại nhằm mục đích tấn công mạng:

- Đánh lừa hệ thống bảo mật: tội phạm phát triển và sử dụng phần mềm tích hợp AI để vượt qua hệ thống bảo mật của thiết bị.

- Phần mềm đánh cắp dữ liệu: Sau khi được cài đặt, các phần mềm đánh cắp dữ liệu yêu cầu người dùng cung cấp quyền truy cập hệ thống và đánh cắp thông tin như hình ảnh, danh bạ, tin nhắn, tài liệu được lưu trữ trên thiết bị nhằm quét dữ liệu hình ảnh, nhận dạng các cụm từ ghi nhớ tiềm năng từ hình ảnh được lưu trữ trên thiết bị và đăng tải dữ liệu định kỳ đến một máy chủ xa nhằm trích xuất thông tin đăng nhập của người dùng, từ đó thực hiện các hành vi tấn công hệ thống an ninh mạng, lừa đảo chiếm đoạt tài sản.

- Phần mềm đoán mật khẩu: các phần mềm sử dụng công nghệ mạng đối nghịch tạo sinh (GANs) phân tích tập dữ liệu lớn về mật khẩu và tạo ra các biến thể phù hợp với phân bổ thống kê nhằm đoán mật khẩu có hiệu quả.

- Phần mềm phá vỡ hệ thống CAPTCHA: CAPTCHA là rào cản đối với các cuộc tấn công mạng tự động.

- Tấn công hệ thống mạng: Tội phạm mạng đã nghiên cứu, phát triển các công cụ để thực hiện các cuộc tấn công hệ thống mạng như: Thực hiện việc "hack" hệ thống mạng không dây (Wifi) thông qua các cuộc tấn công hủy xác thực (de-authentication attacks).

4. Sử dụng AI để phạm tội trên thị trường tiền điện tử:

- Kêu gọi đầu tư vào những dự án lừa đảo và lan truyền trên quy mô lớn: AI tạo sinh và Deepfake được sử dụng để tạo ra những hình ảnh, video quảng cáo xuất hiện người nổi tiếng, có sức ảnh hưởng để tăng uy tín, sau đó đăng tải trên mạng xã hội nhằm kêu gọi người tham gia đầu tư vào các dự án tiền điện tử lừa đảo. Đồng thời, tạo ra nội dung quảng cáo không có thật, sau đó đăng tải lên các diễn đàn tiền điện tử nhằm củng cố niềm tin của nhà đầu tư.

- Tạo ra các đồng tiền điện tử lừa đảo liên quan đến AI và thao túng thị trường: các đối tượng sử dụng công nghệ AI tạo ra các đồng tiền điện tử lừa đảo một cách dễ dàng trên chuỗi khối (blockchain). Đồng thời tạo ra giao dịch ảo hoặc chênh lệch giá, từ đó đẩy giá tiền điện tử để thao túng thị trường.

- Tạo ra các trang web giao dịch lừa đảo và chương trình tặng tiền điện tử: AI cho phéptội phạm thiết kê giao diện trang web giao dịch lừa đảo một cách nhanh chóng và tự động "sập" sau khi tội phạm đã huy động được lượng tiền lớn. Bên cạnh đó, tội phạm sử dụng AI để tạo ra các chương trình tặng tiền điện tử miễn phí cho nhà đầu tư hoặc khuyến khích nhà đầu tư mua tiền điện tử sớm với giá hời và hứa hẹn giá trị tiền điện tử sẽ tăng rất nhiều lần sau khi phát hành trên các sàn giao dịch điện tử để củng cố niềm tin của nhà đầu tư.

- Phát hiện lỗ hổng giao dịch để trục lợi: Lợi dụng việc hầu hết các giao dịch tiền điện tử phi tập trung (Dcfi) sử dụng mã nguồn mở, có thể xem công khai hoạt động giao dịch và lưu trữ tài sản của nhà đầu tư, tội phạm đã sử dụng AI để phát hiện và khai thác các lỗ hỏng giao dịch để đánh cắp tiền điện tử.

Thời gian tới, tình hình tội phạm liên quan đến AI sẽ diễn biến hết sức phức tạp với nhiều thủ đoạn tinh vi và đa dạng trên nhiều lĩnh vực như an ninh – quốc phòng, kinh tế, văn hóa – xã hội, tạo ra những thách thức rất lớn đối với hệ thống pháp luật, nhất là hành lang pháp lý liên quan đến công tác bảo đảm an ninh mạng và phòng, chống tội phạm sử dụng công nghệ cao, vấn đề này đang đối mặt với nhiều khó khăn, phức tập, cụ thể như:

- Tội phạm hoạt động trên phạm vi toàn cầu, không bị giới hạn bởi biên giới quốc gia, gây khó khăn cho việc truy tìm, truy vết và xử lý tội phạm;

- Tội phạm có kỹ thuật tinh vi, bao gồm việc sử dụng mã độc, mã hóa thông tin và thay đổi vị trí truy cập, gây khó khăn cho công tác điều tra;

- Người dân còn thờ ơ về các mối đe dọa từ an ninh mạng, dễ trở thành nạn nhân của các thủ đoạn lừa đảo, tấn công mạng hoặc mã độc; chưa có kỹ năng về bảo mật thông tin cá nhân trên không gian mạng, đặc biệt là các thông tin được lưu trữ trực tuyến như tài khoản ngân hàng, thẻ tín dụng, số điện thoại, địa chỉ liên hệ, email,…

- Công nghệ phát triển nhanh chóng, thường xuyên thay đổi, vượt qua giới hạn kiểm soát của cơ quan chức năng, gây rất nhiều khó khăn cho công tác phát hiện, ngăn chặn và xử lý loại tội phạm mới này;

- Nhiều quốc gia trên thế giới chưa có hiệp định tương trợ tư pháp trong việc xử lý các vụ việc liên quan đến tội phạm AI, trong khi tội phạm hoạt động trên phạm vi toàn cầu, gây khó khăn cho việc áp dụng và thực thi pháp luật.

Ở nước ta, trong thời gian qua, tình hình tội phạm và vi phạm pháp luật liên quan đến AI diễn biến rất phức tạp. Để phòng ngừa, đấu tranh chống tội phạm sử dụng trí tuệ nhân tạo, Sở GD&ĐT đề nghị Thủ trưởng các đơn vị thông

tin tuyên truyền các thủ đoạn trên cho toàn thể công chức, viên chức, người lao động cùa ngành biết và nêu cao tinh thần cảnh giác, không để các đối tượng sử dụng trí tuệ nhân tạo lừa đảo chiếm đoạt tài sản, tránh những hậu quả đáng tiếc xảy ra. Khi phát hiện trường hợp có dấu hiệu nghi vấn, cần báo ngay cho cơ quan Công an nơi gần nhất để kịp thời xử lý./.

Nơi nhận:
- Như trên;
- GĐ, PGĐ Sở;
- Các phòng Sở;
- Thanh tra Sở;
- Lưu: VT, HCQT.

KT. GIÁM ĐỐC
PHÓ GIÁM ĐỐC

SỞ GIÁO DỤC VÀ ĐÀO TẠO

Nguyễn Hồng Phúc