中国は生成 AI を使ったサイバー攻撃を開始、Microsoft は東アジアのセキュリティリスクを分析、日本や米国に対する情報操作の脅威が増すと警告

2024/5/3 研究レポート #1039

Microsoft は AI を使ったサイバー攻撃に関する分析レポート「Microsoft Threat Intelligence」を公開した。これは東アジアにおける脅威を分析したもので、中国は生成 AI など高度な技術を導入し、攻撃手法が進化していると警告。福島原子力発電所の処理水の放出に関し、生成 AI で作成した偽画像が使われ、危機感を煽るキャンペーンが展開された。台湾の総統選挙においては、AI で生成したイメージが急増した。米国では、大統領選挙に向けて、国民世論を分断する試験が繰り返されていると警告している。

出典: Microsoft

サイバー攻撃分析レポート

このレポートは Microsoft のサイバー攻撃分析センタ「Microsoft Threat Analysis Center (MTAC)」が発行したもので、中国と北朝鮮によるサイバー攻撃の実態と動向を分析している。レポートは、サイバー攻撃の特徴として、件数が増大したことに加え、生成 AI が導入され、攻撃技術のレベルが上がったと指摘する。従来からサイバー攻撃に AI が使われているが、生成 AI を導入することで高精度な偽画像を容易に生成できるようになった。

レポートの要旨

レポートは、従来型のサイバー攻撃に加え、ソーシャルメディアを使った情報操作の技術が向上し、危険性が増大したと結論付けている。サイバー攻撃は二種類あり、１）サイバー攻撃(Cyber Operations)と２）情報操作(Influence Operations)となる。前者はマルウェアなどによる従来型のサイバー攻撃で、後者はソーシャルメディアを使った情報操作を指す。レポートの要旨は：

- 中国：南太平洋諸島や南シナ海や米国の軍事企業を対象にしたサイバー攻撃が継続されている。情報操作活動については、生成 AI など新しい技術を導入し、その実証試験を通じ、効果の検証を進めている。
- 北朝鮮：サイバー攻撃が中心で、ソフトウェア・サプライチェーン攻撃やランサムウェア攻撃で重大な被害が発生している。

中国による情報操作

レポートは中国による情報操作を特に警戒している。生成 AI など高度な AI を使い、イメージを生成・編集するもので、これらをソーシャルメディアで拡散し、世論分断などの情報操作を実行する。ビデオやイメージや音声などが使われ、攻撃対象は米国の他に、台湾、日本、韓国など東南アジアの国々が含まれる。現時点では生成 AI を使った情報操作の試験段階であり、様々な手法が試され、その効果を検証していると分析。

情報操作の事例：福島原子力発電所の処理水放出

中国の情報操作はソーシャルメディアにアカウントを設け、ここから偽情報を発信し国民の世論を操作する手法を取る。このオペレーションでは「Storm-1376」というアカウントが使われ、ここから偽情報が発信された。福島第一原子力発電所が処理水を放出したことに関し、日本政府を非難するメッセージが日本語、韓国語、英語で大量に発信された。この情報操作の特徴は生成 AI で作成されたイメージが使用されたことにある(下の写真左側)。また、他のアカウントのコンテンツを再利用したケースもある(中央と右側)。また、韓国に向けて発信された情報操作では、日本政府の措置に反対する運動を喚起するもので、日本と韓国の分断を助長することを目的としている。

出典：Microsoft

情報操作の事例：マウイ島の山火事

ハワイ・マウイ島で 2023 年 8 月、大規模な山火事が発生し、多くの人が犠牲になった。米国で発生した山火事としては過去 100 年で最悪の被害といわれている。上述のアカウント「Storm-1376」は山火事に関して偽情報を複数のソーシャルメディアで発信した。山火事は米国政府が「気象兵器(Weather Weapons)」を試験するために意図的に出火したものであるとの陰謀論を展開。ソーシャルメディアに海岸に面した住宅地での火災の写真が掲載されたが、これらは AI でイメージを誇張したもので、読者の危機感を煽る仕組みとなっている(下の写真)。

出典: Microsoft

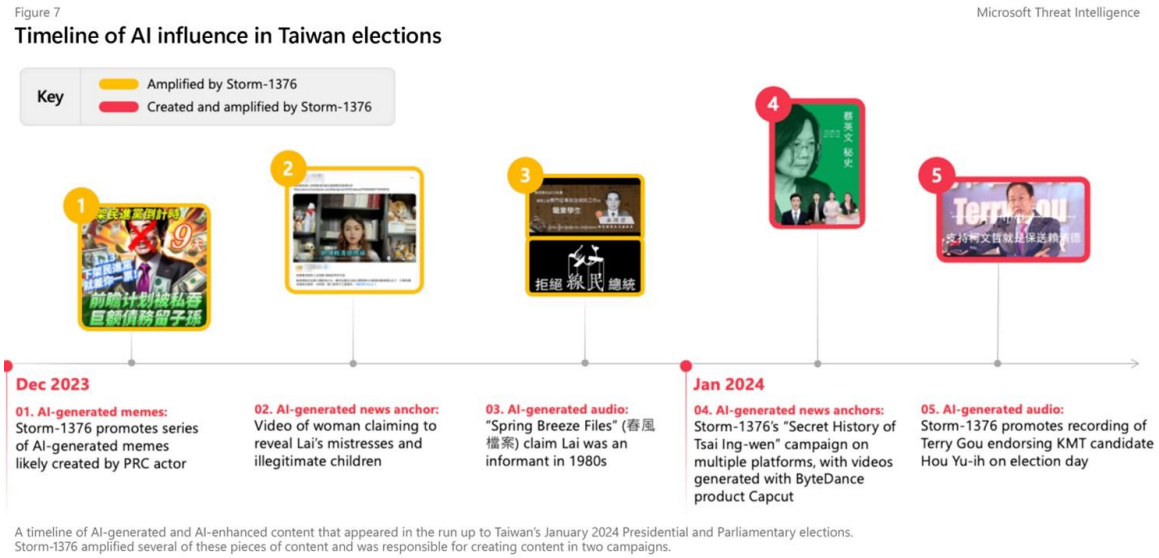
米国大統領選挙に向けた攻撃準備

中国の情報操作は米国においては、大統領選挙に向け攻撃手法の準備を目的に進められている。実際に、米国の有権者の意見を理解するためのオペレーションを開始した。米国で世論が二分されているテーマについて取り上げ、有権者の意見を聴取するコンテンツを発信。地球温暖化、国境警備、違法薬物、移民政策、人種問題に関する写真などを掲載し、有権者に「国境警備の費用に 200 億ドルの予算が充てられるが、これをどう思うか」などと問いかける(下の写真右側)。国民の考え方を把握し、大統領選挙では国民の世論を分断する偽情報を発信することを目的としている。

出典: Microsoft

主要国で選挙が行われる

今年は、インド、韓国、アメリカで重要な選挙が行われる年で、中国はこの機会を利用して世論操作を展開するとレポートは分析している。既に、1 月に実施された台湾の総統選挙では、AI で生成したイメージやボイスが使われ、情報操作の新たな手法が示された(下の写真、コンテンツは AI で誇張したイメージやボイスから、AI で生成したものに進化)。偽のイメージやボイスを合成するために生成 AI が使われており、これらを検知する技術の確立が求められる。



出典: Microsoft

生成 AI による攻撃をどう防ぐか

米国に対するサイバー攻撃はロシアが主導してきたが、ウクライナ戦争の影響なのか、米国における活動が低下している。この空白を埋めるように、今では中国が米国に対する情報操作活動を展開している。攻撃ツールとして生成 AI が使われ、警戒感が高まった。生成 AI による攻撃手法を完全に把握できてなく、これをどう防御するのか議論が広がっている。生成 AI による攻撃は、生成 AI で防御すべきとの考え方もあり、セキュリティ技術の開発が喫緊の課題となる。