New Paths of Attacks: Revealing the Adaptive Integration of Artificial Intelligence in Evolving Cyber Threats Targeting Social Media Users and Their Data

Larry Teasley
Old Dominion University

New Paths of Attacks: Revealing the Adaptive Integration of Artificial Intelligence in Evolving Cyber Threats Targeting Social Media Users and Their Data

By: Larry E. Teasley III

Old Dominion University

Department of Computer Science

Professor Malik A. Gladden

December 1, 2023

Introduction

Many different artificial intelligence (AI) tools are becoming available for society to use for many purposes. This creates opportunities for bad actors to exploit these new tools to attack several vulnerabilities in a social media-dependent society. In recent years, we have seen a significant increase in cyber-attacks on social media platforms, with malicious actors attempting to gather user data and personal information. Several social media companies have tried to reduce the number of attacks on their platform; however, as new technologies continue to develop, social media has become more prone to attacks. Bad actors can gather large amounts of user information on social media platforms through various ways, including generative AI. Malicious actors have used generative AI to enhance phishing attacks on social media and cause other dysfunction and chaos by using this AI to spread false information and propaganda. Malicious actors can also easily access large language models (LLMs) to enhance their fake social media posts by learning from summarized and translated text to resemble how humans communicate. This poses an even more significant challenge for social media companies and users in differentiating between a generative or fake social media account and a real one. Although this may not be an attack on a physical system or server, it is still considered a cyber-attack because of how the attackers orchestrate their attacks by manipulating cyber platforms to target individuals. Throughout this research paper, the many new challenges that target social media platforms and technologies will be revealed, and an explanation as to why they are so detrimental to the public will be given. These challenges will include human dependency on social media, social media user data risks, the implementation of artificial intelligence, enhancing attacks with artificial intelligence, and the advancements of deep fakes.

Human Dependency on Social-Media

Human curiosity and the need for constant advancement are the most reasonable explanations as to why there is a large digital ecosystem that is ever-expanding today. What does not help is that society has become more dependent on new technologies, especially since the COVID-19 pandemic, causing us to live more deeply in the digital world. Social media has become an effective tool for individuals to communicate with others within and outside their communities. Social media allows for the wide spread of information with practically everyone worldwide. Outside of these technologies being used as a form of communication and entertainment, they are also profit centers for the companies that develop them and the social media users, making it more lucrative for individuals to join. With people becoming more intertwined with technology, specifically social media, humanity's best and worst are being enabled. These technologies promote many positives but are also being exploited for various purposes. No matter how convenient of a tool social media is and continues to evolve to be, malicious actors continue to use this effective digital ecosystem to steal data and intellectual property, distribute disinformation, prolificate online harassment, and threaten the peace and stability of different communities around the world, especially in times of conflict (White House, 2023). However, social media companies have continued to market their platforms as a necessity for everyday life, creating a more social media-dependent lifestyle. With this heavy involvement with social media platforms, malicious cyber activity has evolved from nuisance defacement to espionage and intellectual property theft and cyber-enabled influence campaigns designed to undermine public trust in the accuracy of information on current events (White House, 2023).

Social Media User Data Risk

As society becomes more intertwined with different social media technologies, the number and intimacy of user data collection is proliferating. Social media companies, specifically, collect user data to enhance the user experience. Other organizations feel that discussions on social media are essential for collecting data which can be utilized by different machine and deep learning models to identify and predict consumers intentions (Sharma, 2022). However, the cost of doing so creates a grave danger of that data being leaked on a large scale. This was seen recently in a major data leak that occurred on the social media platform Twitter. Twitter has seen increased data leaks recently, but this specific event was responsible for exploiting 5.4 million Twitter accounts (Dent, 2022). The goal behind collecting the user's data is to sell the information to high bidders to carry out other malicious attacks. Some information collected through a vulnerability in Twitter's API was user phone numbers and email addresses (Dent, 2022). With this information, different attacks, e.g., phishing attacks, can be enhanced by making them more indistinguishable because they use real people's information. Cybercriminals are also driven to steal social media users' data and spread disinformation and propaganda during conflict. Cybercriminals will be credible if they can use the data collected from a legitimate social media user to masquerade a hoax account.

User Data Collection

Data collection is understood in three ways: volume, variety, and velocity (Singh, 2020). As described by researchers, data variety includes formats like text, audio, video, and images (Singh, 2020). The evolution of social media platforms has allowed this volume to include those varieties and even more, given the new features of social media applications, like online

shopping. However, collecting all of this data has become easier with traditional methods due to FOMO (fear of missing out), causing more users to be willing to share information and agree to terms that explicitly outline that the social media platforms will store the information shared (Singh, 2020). In addition, people are also loosely sharing information that seems harmless and even encouraged by different social networking platforms; however, that information can be exploited to enhance phishing scams by malicious actors or other fraudulent attacks.

Implementing Artificial Intelligence

A new common trend amongst social media users is using AI for their social media posts. Several companies, including social media companies, have noticed this trend, causing a significant increase in the number of AI-generated social media posts. Malicious actors have also noticed this trend. With easy access to different generative AI, like LLMs, these bad actors have taken advantage of these tools to enhance several different types of attacks. Most cyber-attacks targeting different organizations like government agencies, hospitals, financial institutions, and social media platforms have previously required experienced hackers with knowledge of manual methods to carry out their attacks (Mirsky, 2023). However, today, AI is responsible for tools that enhance automated tasks, making it more efficient and accessible for malicious actors to target these industries. To further show how easy it can be to write programs using AI tools, I utilize ChatGPT to write a program that analyzes social media posts and collects social media user data through a platform's API. Although the program requires additional information for authentication purposes, by utilizing other tools and minimal information gathered through social media platforms like YouTube, a malicious actor can alter the program generated by ChatGPT to fit the needs of a cyber-attack. In today's age, malicious actors no longer need to have an intensive understanding of programming or other computer science topics. If they have access to

the internet and social media platforms, they can easily find and exploit emerging trends and tools that make it easier for them to carry out cyberattacks.

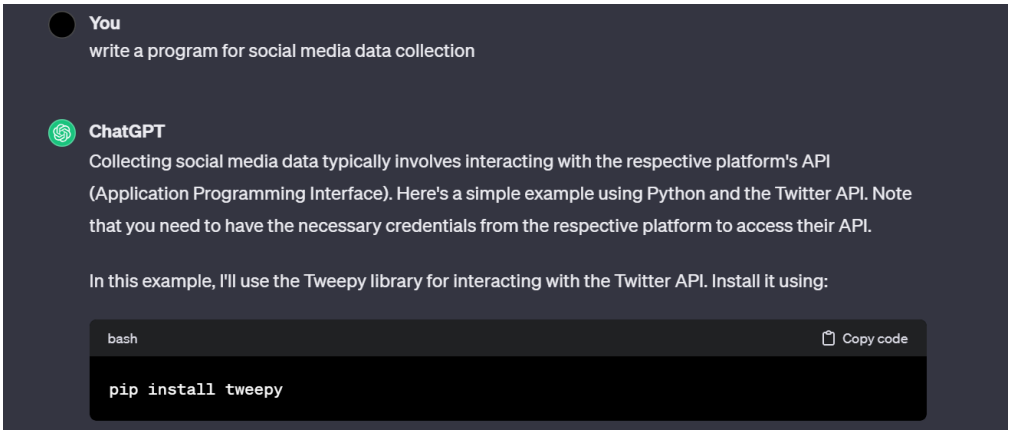

Figure 1: shows the program generated by Chat-GPT and the question it was prompted.

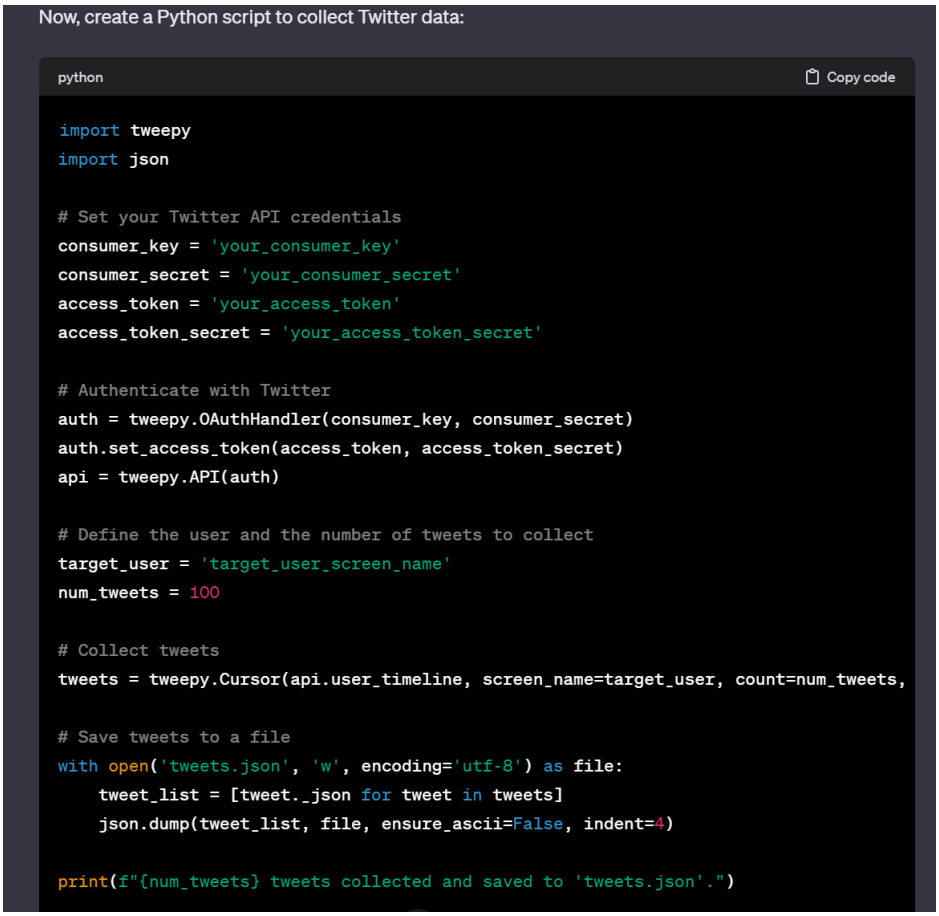

Figure 2: Shows the generated Python program for collecting Twitter data.

Enhancing Attacks with Artificial Intelligence

In the case study conducted by Yamin, new AI-driven attacks were investigated, showing the implementation of AI technologies to the classical cyber security attacks (Yamin, 2021). AI-based techniques will further complicate the job of cyber security professionals, making it harder to detect several of these new attacks that utilize AI. With the current wave of attacks, malicious actors are implementing these technologies to direct targeted attacks at unprecedented speed and scale while avoiding traditional detection measures (Guembe, 2022). With the efficiency of these tools, cybercriminals will be able to socially engineer targets at human or superhuman levels of performance, all while learning from and adapting to the environment in which it is placed (Guembe, 2022). Another study, which outlines the vast dangers of AI implementation in cyberattacks, illustrated how present AI techniques are in the cybersecurity kill chain in the access and penetration phase (Guembe, 2022). The researchers used generative adversarial network (GAN) and recurrent neural network (RNN) techniques to illustrate access and penetration attacks. These researchers also demonstrated the presence of AI techniques in the reconnaissance phase of the cyber security kill chain by utilizing vulnerability prediction, End-to-End (E2E) spear-phishing, and intelligent profiling/ intelligence collection attack methods (Guembe, 2022). In the study, Markov chains and LTSM were used to create automated machine-generated content disinformation for vulnerability prediction, and the E2E spear-phishing technique was used to generate personalized content for high-target users on Twitter (Guembe, 2022). By using these AI techniques for different cyber-attacks to improve reconnaissance, the researchers discovered four AI-driven threats: intelligent target profiling,

clever vulnerability detection/intelligent malware, intelligent collection/automated learn behavior, and intelligent vulnerability/outcome prediction (Guembe, 2022).

Advancements Of Deep Fakes

Also, by implementing AI, hackers can more efficiently create deep fakes to enhance several attacks for exploiting data. A deepfake is a form of AI that generates highly realistic synthetic media, including texts, sounds, videos, and images, by utilizing deep learning technology (NSA, 2023). The advancement of deepfake technology and its implementation into social media allows malicious actors to cause disruption by enhancing social engineering techniques to get people to believe someone with credibility has said or done something they have not. Malicious actors can take this further by creating other fake accounts and fake posts on a social media platform that collaborates with the original deep fake post. On a large scale, it can cause massive chaos in society, potentially destabilizing many industries. This is important because most of this can be done with a piece of software that is practically plug-and-play for an end user, requiring very little to no knowledge of deep fake software. A cybersecurity information sheet published in September of 2023 through the collaboration of the NSA, FBI, and CISA further provides an overview of the different synthetic media cybersecurity threats (NSA, 2023). Although technology to manipulate authentic media has been around for many years, the complexity of leveraging manipulated media has fallen, making it more enticing for cybercriminals with little to no technical expertise to use deepfakes (NSA, 2023). These prolific and relatively available tools enable fraud and disinformation to exploit targeted individuals and organizations (NSA, 2023). Although there has been a dramatic increase in collaborative work on creating and improving detection software for deep fakes, the same has been done to improve deep fake software. Another emerging trend in the development of synthetic media has been the

improved ability to lift a 2D image to a 3D to enable the realistic generation of video based on a single image (NSA, 2023). This will require even more work for organizations and cybersecurity professionals to combat the malicious use of these emerging technologies.

Conclusion

Throughout this research, many challenges that result from the intersection of artificial intelligence with social media, causing new paths of attacks, have been revealed. Social media has been one of the most excellent technical tools to be developed in the 21st century. It has allowed us to obtain and share information around the world and connect with people worldwide in ways that were unimaginable in the 19th century. This tool has solved many problems that society has faced during times of distress and conflict and has allowed society to evolve for the best and the worst. Social media has shaped many industries, including education, health care, and government organizations. Social media has also shaped society by changing social norms and beliefs while providing an outlet for the seamless share of opinions and factual information between individuals worldwide.

Although this excellent resource has many positives, it is exploited to bring out the very worst of humanity. Malicious actors have begun taking advantage of social media platforms to carry out new waves of cybersecurity attacks that target not only the social media platform but also high-value individuals on that platform. This recent evolution of attacks has revealed several challenges that cybersecurity professionals now have to face: human dependency on social media, user data risks, the implementation of artificial intelligence, enhanced attacks with artificial intelligence, and the advancements of deep fakes. Although most of these challenges are not new, how cybercriminals exploit them is. Most malicious actors create new paths of attacks that target the vulnerabilities in our social media-dependent society by taking the

traditional challenges and enhancing them using tools generated by artificial intelligence software.

Exploring the human dependency on social media has illustrated the dichotomy between the tool's positive contributions to society and its exploitation for nefarious purposes. What used to be a harmless outlet for entertainment that carried basic cyber security risk has become a feeding ground for cybercriminals to gather individuals' information, spread false information, and disrupt societies worldwide. The escalated risks associated with social media user data, highlighted by recent data leaks, exemplify the urgency of addressing privacy concerns and implementing robust security measures. As demonstrated by the generated Python program, integrating AI in cyberattacks amplifies its efficiency and ease, even for individuals with minimal expertise in programming, which poses a formidable challenge.

Moreover, with the fact that most people ingest most of their knowledge from social media platforms, with little care to verify the captivating headline that they see on their timeline, malicious actors see this as an opportunity for the creation of bot social media accounts that masquerade as a reliable media outlet for news and other sources of information. They do so by leveraging AI tools, like LLMs, to collect and analyze information from social media platforms to better impersonate a human account, evading the detection of different security measures that social media platforms have in place to detect fake accounts and social media posts. Malicious actors are also leveraging deepfake technology to enhance their attacks further. Given that deep fake technologies are freely available to the public and require very little to no education to use them to their full potential, the increase of cyber-attacks on social media platforms involving deep fakes will continue to grow rapidly and require new solutions to detect them.

Navigating this complex terrain will require the collaboration of social media companies, cyber security professionals, policymakers, and other organizations to create proactive security measures. Just as technical considerations, ethical considerations must guide the development and implementation of AI in society. With better-enforced security protocols to safeguard user data and by taking decisive action to address the evolving risks, there will be a social media landscape and digital ecosystem that fosters connectivity, all while not compromising security protocols and the trust of its users. Nevertheless, the synergistic relationship between social media and AI requires continuous adaptation and collaboration to mitigate the emerging threats of the new paths of attacks.

References

Dent, S. (2022, November 28). *Twitter data leak exposes over 5.4 million accounts*. Engadget. https://www.engadget.com/twitter-data-for-54-million-users-leaked-online-095040426.html?guccounter=1

Guembe, B., Azeta, A., Misra, S., Osamor, V. C., Fernandez-Sanz, L., & Pospelova, V. (2022). The emerging threat of AI-Driven Cyber Attacks: A Review. *Applied Artificial Intelligence*, *36*(1). https://doi.org/10.1080/08839514.2022.2037254

Mirsky, Y., Demontis, A., Kotak, J., Shankar, R., Gelei, D., Yang, L., Zhang, X., Pintor, M., Lee, W., Elovici, Y., & Biggio, B. (2023). The Threat of Offensive AI to Organizations. *Computers & Security*, *124*, 103006. https://doi.org/10.1016/j.cose.2022.103006

National Security Agency, Federal Bureau of Investigation, & Cybersecurity and Infrastructure Security Agency. (2023, September). Contextualizing deepfake threats to organizations. https://media.defense.gov/2023/Sep/12/2003298925/-1/-1/0/CSI-DEEPFAKE-THREATS.PDF

Sharma, A., & Shafiq, M. O. (2022). A comprehensive artificial intelligence based user intention assessment model from online reviews and social media. *Applied Artificial Intelligence*, *36*(1). https://doi.org/10.1080/08839514.2021.2014193

Singh, M., Verma, C., & Juneja, P. (2020). Social Media Security Threats Investigation and Mitigation Methods: A preliminary review. *Journal of Physics: Conference Series*, *1706*(1), 012142. https://doi.org/10.1088/1742-6596/1706/1/012142

The White House. (2023, March). National Cybersecurity Strategy - The White House. https://www.whitehouse.gov/wp-content/uploads/2023/03/National-Cybersecurity-Strategy-2023.pdf

Yamin, M. M., Ullah, M., Ullah, H., & Katt, B. (2021). Weaponized AI for cyber attacks. *Journal of Information Security and Applications, 57*, 102722. https://doi.org/10.1016/j.jisa.2020.102722