MEMORANDUM
TO: Chief Information Security Officer (CISO)
FROM: Kanchan Chopde
DATE: 9th April 2024
SUBJECT: Potential Cybersecurity Vulnerability and Recommended Actions

PURPOSE

The purpose of this memo is to inform you of a potential cyber security vulnerability that has been identified within our organization's systems, and to provide recommendations on the steps we should take to address this issue.

Generative AI-Powered Phishing Attacks

Our organization has experienced Generative AI-Powered Phishing Attacks. Cybercriminals are leveraging advanced language models like Chat GPT to craft highly convincing phishing emails and messages. These AI-generated communications can bypass traditional security filters and trick users into disclosing sensitive information or downloading malware.
Phishing attacks have become one of the most widespread and damaging cyber security threats, with 92% of organizations falling victim in the last 12 months. The use of generative AI to automate and scale these attacks makes them even more dangerous, as the messages appear highly personalized and legitimate.

The ways by which generative AI-powered phishing attacks has spread and infected our systems or users are:

1. Fake Emails and Messages:
   - Cybercriminals are using generative AI models like ChatGPT to create highly convincing phishing emails and messages that mimic the writing style and tone of trusted contacts.
   - These AI-generated communications can bypass traditional security filters and trick users into disclosing sensitive information or downloading malware by including malicious links or attachments.
2. Deepfake Voice and Video:
   - Attackers have leveraged generative AI to create deepfake audio and video content, such as cloning the voice of a trusted contact to make vishing (voice phishing) attacks more convincing.
   - Users may be tricked into providing sensitive information or taking harmful actions if they believe the communication is from a legitimate source. 13
3. Automated Phishing Campaigns:
   - Generative AI has allowed cybercriminals to automate the creation and distribution of phishing attacks at a much faster rate, increasing the overall attack surface.
   - The AI-generated phishing content has been sent to a large number of targets from our organization simultaneously, making these campaigns more scalable and effective.

4. Personalized Social Engineering:

    - Generative AI can quickly gather and curate personal information about our people from various online sources to craft highly targeted and personalized phishing attacks.
    - This level of personalization makes the phishing attempts more believable and likely to trick users into disclosing sensitive data or downloading malware.

Threat Relevance:

1. Application Domain:

    - Generative AI represents an advanced application-level technology that can be weaponized by threat actors to automate and enhance phishing attacks.
    - Cybercriminals are leveraging generative AI tools like language models to craft more convincing and scalable phishing communications.
    - The application domain is directly targeted as the point of exploitation for these AI-powered phishing attacks.

2. User Domain:

    - Generative AI-powered phishing attacks primarily target the user domain by exploiting human vulnerabilities and tricking end-users into disclosing sensitive information or downloading malware.
    - The highly convincing and personalized nature of these AI-generated phishing messages makes them more likely to bypass user defenses.

3. Data Domain:

    - Successful phishing attacks enabled by generative AI can lead to data breaches, exposing sensitive information within the targeted organization.
    - The data domain is at risk of compromise if users fall victim to these advanced phishing tactics.

4. Logical Domain:

    - Generative AI-powered phishing attacks can provide threat actors with initial access to the target's systems, potentially compromising the logical domain.
    - Once inside the network, the attackers can then move laterally and escalate privileges, further endangering the logical domain.

5. Physical Domain:

    - While not directly targeted, the physical domain can be impacted if the generative AI-powered phishing attacks lead to ransomware deployment or other malware that disrupts physical infrastructure and operations.

The repercussions of not addressing the issue of Generative AI-powered phishing attacks are severe:

1. Phishing attacks using AI-generated content will become much more successful and widespread. Cybercriminals are already leveraging tools like ChatGPT to craft highly convincing phishing emails that are harder for users to detect.13 This will lead to a significant increase in successful phishing attempts, with an estimated 1,265% surge in malicious phishing emails.
2. The rise of AI-powered phishing will result in substantial financial losses. Business email compromise (BEC) attacks alone cost over $2.7 billion in 2022, and other phishing attacks resulted in $52 million in losses.These figures are expected to grow as AI makes phishing more effective.
3. Sensitive data and intellectual property will be at greater risk of theft and compromise. Generative AI can be used to create highly targeted phishing attacks that trick employees into disclosing login credentials or other sensitive information.
4. Cybersecurity teams will face significant challenges in detecting and mitigating these AI-powered threats. The sophistication of AI-generated phishing content makes it harder for traditional security tools and user awareness training to be effective.
5. Organizations that fail to properly secure against AI-powered attacks may suffer reputational damage, operational disruptions, and other serious consequences from successful breaches.45

Remediation steps include:

Implement Enhanced User Awareness Training:

- Provide regular security awareness training and simulated phishing exercises.
  Estimated Cost: 40 hours x $250/hour = $10,000

Deploy Advanced Email Security Controls:

- Implement AI-powered email security solutions to detect and block sophisticated phishing attempts.
- Integrate with threat intelligence feeds to stay up-to-date on the latest attack vectors.
  Estimated Cost: 80 hours x $250/hour = $20,000.

Conduct Vulnerability Assessments and Penetration Testing:

- Perform regular assessments to identify potential weaknesses in the organization's security posture.
- Engage a third-party security firm to conduct comprehensive testing and provide recommendations.
  Estimated Cost: 120 hours x $250/hour = $30,000

Strengthen Incident Response and Threat Hunting Capabilities:

- Enhance the organization's ability to detect, investigate, and respond to security incidents.
- Implement advanced threat hunting techniques to proactively identify and mitigate emerging threats.
  Estimated Cost: 100 hours x $250/hour = $25,000

Conclusion:

In conclusion, the failure to address the growing threat of Generative AI-powered phishing attacks could expose our organization to a wave of increasingly successful and damaging cyberattacks, resulting in substantial financial losses, data breaches, and other severe repercussions. Proactive measures to enhance security controls and user awareness are critical to mitigating this emerging risk

Please let me know, if any further information is required from my side.

REFERENCES:

1)Embracing Generative AI in Cybersecurity: A Guide for Professionals, Decision-Makers, and Developers (ust.com)
2)Using Generative AI and MongoDB to Tackle Cybersecurity’s Biggest Challenges | MongoDB
3)Generative AI is making phishing attacks more dangerous | TechTarget
4)What Generative AI Means for Cybersecurity in 2024 | Trend Micro (US)
5) Generative AI: Revolutionizing Cybersecurity & Hacker Methods (tokenring.com)
6)AI like ChatGPT is creating huge increase in malicious phishing email (cnbc.com)
7)https://assets.barracuda.com/assets/docs/dms/barracuda-cybernomics-report.pdf

Sincerely,
Kanchan Chopde