VISION

Threat Modeling

Seran Sathyaseelan
April 10,2024

Table of Contents

1.1
Executive summary

As part of our ongoing commitment to ensuring the security and integrity of VISION's AI system, a comprehensive STRIDE threat model analysis was conducted. This report presents the findings of this analysis, which identified potential security threats across various components of the system, including the client's cloud server, data flow, computers, and web application. By understanding and addressing these threats, VISION can strengthen the security posture of its AI system, safeguard sensitive data, and maintain the trust and confidence of its clients and stakeholders.

The STRIDE Threat Dragon Model analysis for the system architecture of VISION reveals several critical security threats and corresponding mitigation strategies. Firstly, the Client's Cloud Server (Store) faces a significant risk of repudiation, where clients may deny their actions, potentially leading to disputes over data integrity.. Additionally, within VISION's Cloud (Store), the threat of tampering poses a considerable risk due to potential model poisoning attacks resulting from insecure coding practices or inadequate monitoring of model activity. Moreover, the transfer of captured images and analysis results faces the threat of information disclosure during transmission and communication, respectively, emphasizing the need for strong authentication mechanisms, access controls, encryption protocols, and regular auditing of access logs to ensure data confidentiality. Finally, both Computers and the Web Application processes are vulnerable to spoofing attacks, necessitating the implementation of robust authentication mechanisms such as multi-factor authentication (MFA), user session management, and secure credential storage to prevent unauthorized access and identity spoofing. By diligently addressing these identified threats through the recommended mitigations, the system can bolster its security posture, safeguard sensitive data, and uphold the integrity and confidentiality of its operations.

2.1
Threat model Overview

The STRIDE threat model is a framework used to identify and categorize potential security threats in software systems. It was developed by Microsoft and is widely adopted in the field of cybersecurity for its structured approach to analyzing and mitigating security risks. STRIDE stands for Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privilege, representing different categories of threats that can compromise the confidentiality, integrity, and availability of a system.

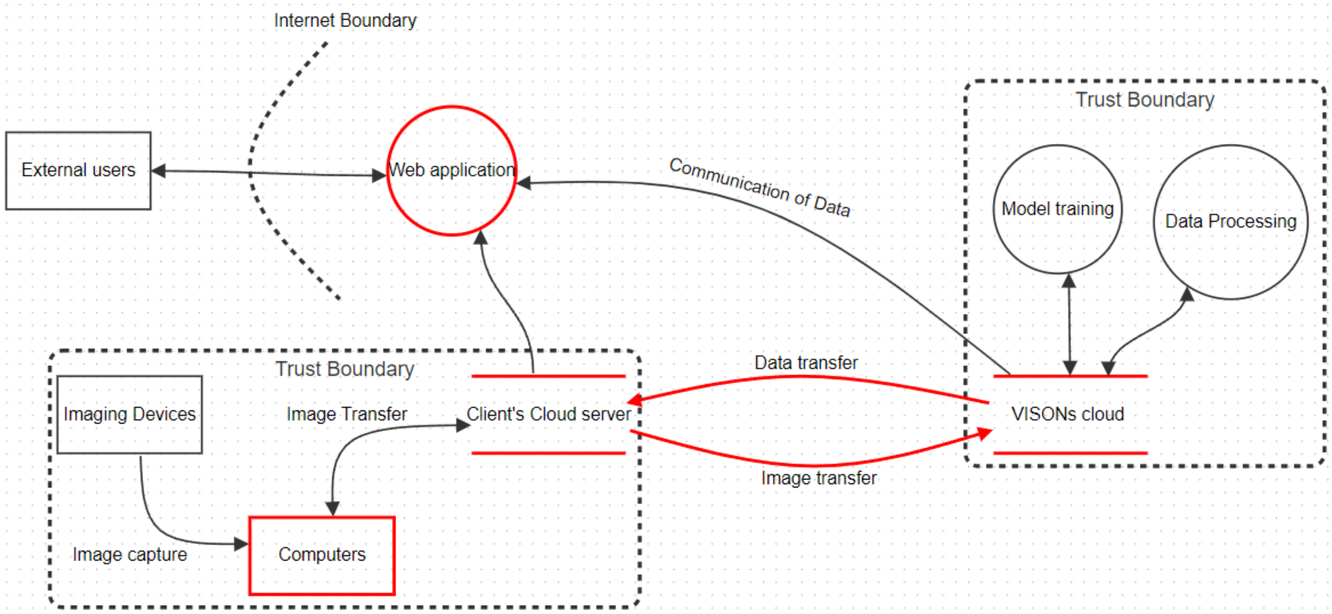
Each category of threat in the STRIDE model represents a specific type of attack vector:

- Spoofing: Involves impersonating or masquerading as legitimate users or systems to gain unauthorized access.
- Tampering: Involves unauthorized modification or alteration of data or systems.
- Repudiation: Involves users denying their actions or transactions within the system, leading to disputes over data integrity.
- Information Disclosure: Involves unauthorized access to sensitive information.
- Denial of Service: Involves attackers disrupting the availability or functionality of a system.
- Elevation of Privilege: Involves attackers gaining higher levels of access or privileges within a system.

Using the STRIDE threat model in VISION allows for a systematic analysis of potential security threats to its AI system. By categorizing threats into these six categories, VISION can identify and prioritize security risks, develop appropriate mitigation strategies, and implement security controls to protect its systems, data, and operations. Additionally, the STRIDE model provides a common language and framework for communication among stakeholders, enabling effective collaboration and decision-making in addressing security concerns. Overall, leveraging the STRIDE threat model helps VISION enhance the security posture of its AI system and mitigate potential risks effectively.

2.2
STRIDE Model Analysis

Figure 1.1 - Threat Dragon STRIDE model



Starting at the beginning of this model, clients data are captured with various imaging devices such as MRIs, CT scans and X-ray machines which are then fed to the computers and stored temporarily before uploading to the clients cloud infrastructure. From the clients cloud infrastructure there is transfer of images to VISIONs cloud infrastructure where the AI model training and data processing takes place. From here the interpreted data are communicated to external users such as healthcare professionals outside clients company, insurance companies and patients. In some instances interpreted data is also shared to clients as well for review before sending to external parties.

2.3
Mitigation and Controls

Based on the threat model certain compensating controls that can be implemented, along with its mitigation strategies are as follows:

Client's Cloud Server

Threat: Repudiation
Control Type: Administrative Control
Mitigation Strategies:
Implement Robust Logging and Auditing Mechanisms: Deploy logging and auditing mechanisms that capture detailed information about user activities, system events, and data access. Ensure that logs include timestamps, user identifiers, and actions performed. Use tamper-evident logging techniques to detect and prevent unauthorized modifications to log data.
Apply digital signatures or cryptographic techniques to log entries to ensure their integrity and authenticity. This helps prevent the manipulation or tampering of log data, enabling reliable auditing and accountability. Educate users about the importance of accountability and the consequences of repudiation. Provide training on how to use systems responsibly, adhere to organizational policies, and understand the implications of their actions. Promote a culture of transparency and accountability within the organization.

VISION's Cloud

Threat: Tampering
Control Type: Technical Controls
Mitigation Strategies:
Implement Input Validation and Filtering Mechanisms: Develop and deploy input validation and filtering mechanisms to sanitize and validate incoming data inputs. Validate data formats, lengths, and content against predefined criteria to prevent injection attacks, buffer overflows, and other common tampering techniques. Establish code review processes to identify and remediate vulnerabilities in software code that could be exploited for tampering. Enforce secure coding practices, such as input validation, output encoding, and parameterized queries, to minimize the risk of code injection and manipulation. Deploy RASP solutions that monitor application behavior and detect anomalous activities indicative of tampering attempts. RASP tools can dynamically apply security controls, such as runtime code integrity checks and behavioral analysis, to prevent tampering in real-time.

Data Flow (Transfer of Captured Images)

Threat: Information Disclosure
Control Type: Technical Controls
Mitigation Strategies:
Strong Authentication Mechanisms and Encryption Protocols: Implement strong authentication mechanisms, such as multi-factor authentication (MFA), to verify the identities of parties involved in data transmission. Utilize encryption protocols, such as SSL/TLS, to secure data in transit and protect it from interception or eavesdropping. Enforce access controls to restrict access to sensitive data during transmission. Implement role-based access controls (RBAC) and least privilege principles to ensure that only authorized entities can access and transmit sensitive information. Monitor and audit access logs regularly to detect and respond to unauthorized access attempts or breaches promptly. Review access logs for unusual activities, unauthorized access attempts, or deviations from normal behavior, and investigate any anomalies detected.

Data Flow (Transfer of Analysis and Interpretation Results)

Threat: Information Disclosure
Control Type: Technical Controls
Mitigation Strategies:
Strong Encryption Protocols: Utilize strong encryption protocols, such as Advanced Encryption Standard (AES), to encrypt analysis results during transmission. Encrypt data at rest and in transit to protect it from unauthorized interception or disclosure. Implement secure communication channels, such as virtual private networks (VPNs) or encrypted tunnels, to prevent unauthorized interception or tampering of data during transmission. Use secure protocols and cryptographic algorithms to establish secure connections and protect data integrity. Conduct regular security assessments and penetration tests to identify and remediate vulnerabilities in communication channels and protocols. Assess the effectiveness of encryption mechanisms, communication protocols, and security controls to ensure robust protection of data during transmission.

Computers (Actor) and Web Application (Process)

Threat: Spoofing
Control Type: Technical Controls
Mitigation Strategies:
Strong Authentication Mechanisms: Implement strong authentication mechanisms, such as multi-factor authentication (MFA), to verify the identities of users and prevent unauthorized access. Require users to provide multiple forms of authentication, such as passwords, biometrics, or security tokens, to access systems or applications. Store user credentials securely using encryption and hashing techniques to protect them from theft or compromise. Use salted hashing algorithms to hash passwords before storage and employ secure key

management practices to safeguard encryption keys. Implement session management controls to detect and prevent session hijacking or impersonation attacks. Use secure session tokens, session timeouts, and session revocation mechanisms to manage user sessions securely and mitigate the risk of unauthorized access.

By implementing these expanded mitigation strategies, VISION can establish a comprehensive security posture to address the identified threats effectively and safeguard its systems, data, and processes from potential risks and vulnerabilities. Regular monitoring, testing, and updates are essential to ensure the continued effectiveness of these mitigation measures against evolving threats and emerging attack vectors.

Closing Remarks

In conclusion, threat modeling serves as a vital tool for identifying, analyzing, and mitigating cybersecurity risks within organizations like VISION. By systematically evaluating potential threats and vulnerabilities across various components of their ecosystem, VISION can proactively anticipate and address security challenges before they escalate into significant incidents.

Threat modeling not only enhances the organization's security posture but also fosters a culture of proactive risk management and resilience. By engaging stakeholders, including technical teams, management, and even clients, in the threat modeling process, VISION can leverage collective expertise and insights to develop comprehensive and effective mitigation strategies.

Furthermore, threat modeling is not a one-time exercise but rather an iterative process that should be continuously refined and updated to adapt to evolving threats and changes in the technological landscape. Regular reviews, assessments, and updates ensure that VISION remains vigilant and responsive to emerging cybersecurity risks.

Ultimately, by integrating threat modeling into its cybersecurity practices, VISION can enhance its ability to anticipate, prevent, and mitigate potential security incidents, thereby safeguarding its systems, data, and reputation in an increasingly digital and interconnected world.