Big Data in Cybersecurity

Dhruval Patel

Eunice Olorunshola

What is Cybersecurity and why is it important?

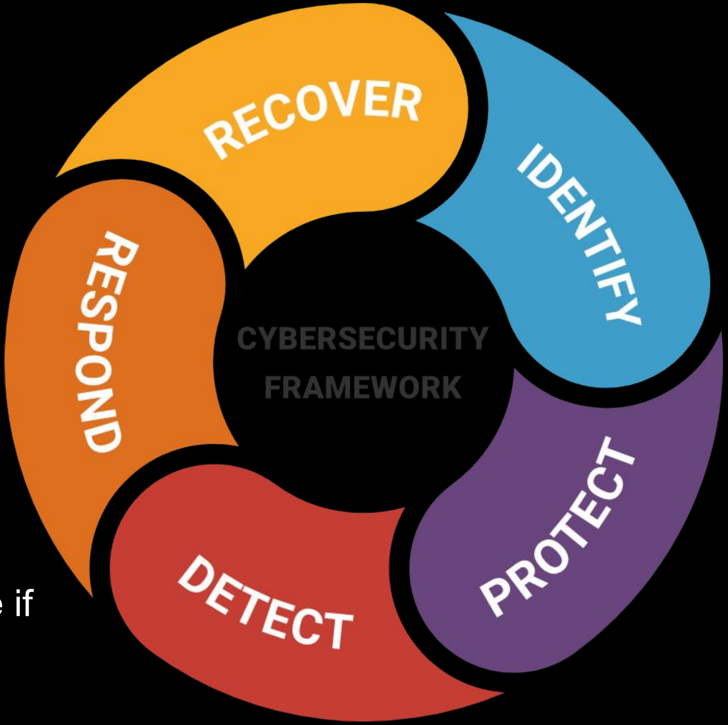
- Protecting devices, networks, data and many more from unauthorized access and criminals.
- **Main goal**: Prevent threats, spams and wrong usage of the data.
- So why is Cybersecurity so important? Why do you hear this phrase so much?
- Everyone relies on technology today
  - Communication
  - Entertainment
  - Shopping
  - Managing confidential information
  - *List goes on*
- Protects all categories of data from theft and damage
  - Sensitive data
  - Personally identifiable information (Personal info)

Why should we use Big Data in Cybersecurity?

- Help us in improving the detection of unusual activity of users, accounts, system, etc.
- Keeps track of historical data from previous cyber attacks in order to discover attack patterns and devise a solution to avoid the same attack from happening again.
- Analyze real time malicious activity that happens in the system
  - Example: Track overall health of the systems by monitoring the data from proxy logs.

How can we use Big Data in Cybersecurity?

- NIST cybersecurity framework
  - Identify
  - Protect
  - Detect
  - Respond
  - Recover
- Real life example: Securing your home.
  - Identify: A lot of robberies in the area
  - Protect: Install security system, camera system,etc.
  - Detect: If someone breaks in then sensors will get triggered
  - Respond: Security system will alert the authorities
  - Recover: Catch the robber and recover any damage if there is any.
- Exact same concept in Big Data.



Example #1: Credit card fraud detection

- Couple different ways frauds are detected:
    - Monitors user's transaction amounts, location, device being used, IP address, time, etc.
    - Multiple purchases made on the same day that are large and do not match the user's transaction activities are flagged.
    - Companies keeps track of a heat map where there is consistent fraud from a specific merchant and notifies the user immediately once a fraud is detected.
- Alerts the user through text message in order to verify if the transaction was authorized or not.

Example #2: Account access protection

- The software collects data by monitoring the user behavior when it comes to the type of device, user interface interaction, IP address, location, time, operating systems, etc.
- The gathered data is used as a user profile to ensure that the systems can detect when there is a suspicious activity and blocks it before an attack occurs.
- When the system detects suspicious behavior, it alerts the user to provide additional authentication.
- Example: When a student logs in their email on GSU’s computer.

What tools/products are used in Big Data Cybersecurity

- Crowdstrike
- Splunk
  - Splunk SOAR (Previously known as Splunk Phantom)
  - Splunk SIEM (Security information and event management)
- Cybereason
- IBM Security
- LogRhythm
- RSA
- *Many more*

Crowdstrike

- Leader in cloud-delivered, next-generation services for endpoint protection, threat intelligence, and response.
- Threat Graph:
    - Collects more than 400 different types of endpoint behavior, spanning Windows, Linux, and macOS, from both user space and kernel space.
    - Actively and automatically enriches and processes the data to reveal and block the most relevant threats in real time.
    - Provides analysts and integrators with real-time, forensic-level visibility into all endpoint activity, no matter how large the organization or how complex the query.

Splunk SOAR

- SOAR is one of the cybersecurity analytics tool and it stands for security orchestration, automation, and response.
- Includes components security automation and orchestration.
- Gives the security administrators easy access.
- Helps to automatically and easily identify the type of attack.
- Security teams are able to utilize more time.

Splunk SIEM

- SIEM is another cybersecurity analytics and it stands for security information and event management.
- SIEM tool overview.
- Includes security concepts SEM(security event management) and SIM(security information management).
- Identifies potential threats from different data sources.
- Role in the security operations center.

Recent study: A Routine Activities Approach to Evidence-Based Risk Assessment

- Purpose: To assess efficacy of the routine activity theory (RAT) for explaining phishing victimization and guide evidence-based policy.
- Simulation: Two phishing emails were sent to both employees and students (total of 25,876 participants).
    - Email 1 contained an embedded link with access to a pdf
    - Email 2 contained an embedded link for free concert tickets

Recent study: A Routine Activities Approach to Evidence-Based Risk Assessment (Cont.)

- Findings:
  - One of the two email attacks sent out, students were less likely to open the email containing the pdf
  - Students were more likely to open the embedded “free concert ticket” link versus the embedded pdf link in the phishing attack emails.
  - Students were more likely than university employees to access the phishing email and embedded link by using their mobile devices
  - When assessing the risk of the University’s network, employees showed more likely than students to open the phishing email when connected to the university network.
  - Out of the two phishing attacks, the email containing the link to the free concert tickets were clicked 2x more than the email containing the link to the pdf file.

Recent study: A Routine Activities Approach to Evidence-Based Risk Assessment (Cont.)

- Conclusion: Applying criminology theory coupled with computer science to analyze information security behaviors, criminology theory was found to be effective in predicting the behavioral patterns in the cyber security environment.

Data Breach

- Data breach is when protected information is exposed to an unauthorized organization or individual. And it gives access of the flies exposed to be viewed and shared without permission.
- The reason why data breach is so important is because most of the information in data breach is confidential which causes the company, organization or an individual to lose a lot of money and many more.
- Prevention of data breach
- The challenges it causes for big data in cybersecurity
- Methods hackers uses to make it happen
- Reasons data breach occurs