Parshvanath Charitable Trust's
A. P. SHAH INSTITUTE OF TECHNOLOGY, THANE
(All Programs Accredited by NBA)
Department of Information Technology

AI Based Soar

Group No. 12

Rahul Vast : 17104042
Shruti Sawant : 18204001
Aishwarya Thorbole : 18204002

Project Guide : Prof. Vishal Badgujar

Contents

Introduction

- Soar stands for Security Orchestration, Automation and Response
- Our project AI Based Soar, main purpose is to gather security related data in real time, perform analysis of that data and make a decision based on that data.
- The components of our system will be SIEM, Automated STIX Engine, Translated Multilingual Threat Data, Threat Repository and Honeypot.
- The Security orchestration part is performed by the components of our system.
- The Automation part is performed using various opensource tools like ELK, Spiderfoot, etc.
- The Response part is handled by our AI’s decision making skills using Neural Network.

Objectives

- To translate a multilingual threat intelligence.
- To recognize different patterns from the event logs/dataset, detection of the vulnerabilities in the system and patch them .
- The data collected from different sources will be converted into a standardized format based via Automated STIX Engine for Privacy Preservation.
- Event profiling of the data is to be done , so there would an ease to find the intelligence of previously known attack.
- To efficiently differentiate between false positive and true positive types of alerts.
- To identify almost all form of attacks and respond to types of attacks like Malware Attacks, SQL Injection, MITM, Password Attacks and DDOS.

Problem Definition

- Traditional Methods for Threat Detection have the inability to gather appropriate/relevant data to be analysed for true positives. Also these methods lack quick response mechanisms for preventing modern attacks.
- IDS may not be potentially always correct, it may include much more false positives. Here a new threat intelligence technique evaluated by analysing honeypot data to identify attackers behaviour to find attack patterns.
- The multilingual nature of the Internet increases complications in the cybersecurity community’s ongoing efforts to strategically mine threat intelligence , from sources such as social media blogs and dark web vulnerability markets that exists in the diverse languages that security analysts often hinder about.

Technology Stack

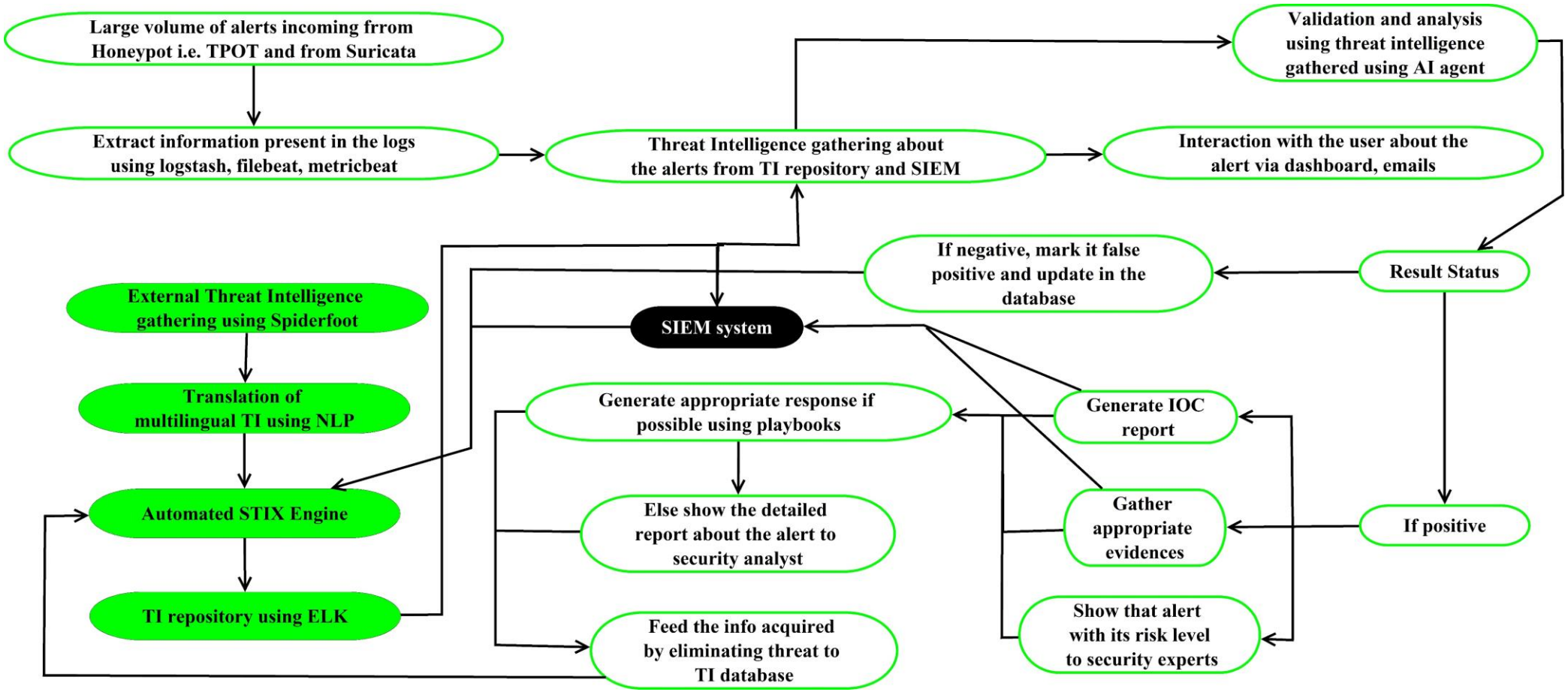
- Python for programming the business logic
- Autokeras for utilising automl capabilities
- Tensorflow for developing AI agent
- Pandas for data pre-processing
- AWS for computational power, storage and deployment
- ELK Stack for data visuslizations and analysis
- Red hat Ansible for automating configurations
- Docker for creating container and ease deployment
- Spiderfoot for automated OSINT gathering
- p0f for passive TCP/IP stack fingerprinting
- Suricata for Network Security Monitoring Engine
- Nginx for reverse proxy

- Cyberchef for encryption, decryption, encoding, etc.
- Cockpit for server administration
- Tpotce for honeypot
- OWASP Tools for activities like vulnerability assessment, application security, testing, etc
- Fatt for network meta data extraction
- Cisco Packet Tracer

Review Suggestions

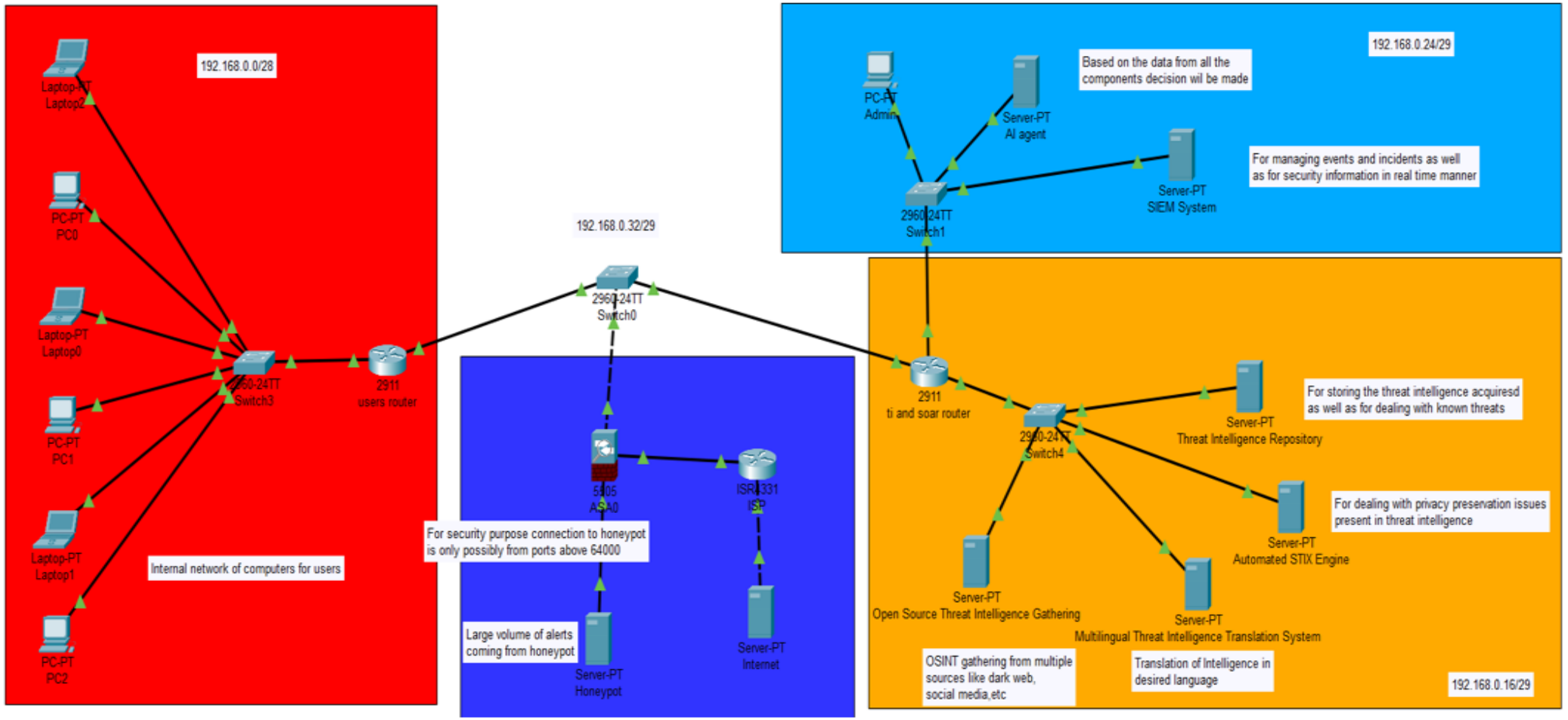
- Be specific with the objectives.
- Highlight the multilingual threat intelligence in objectives.
- Represent the siem diagram using Dia tool.
- Include the types of attacks after the objectives.
- Mention the tools and techniques in proposed system.

Proposed System Architecture/Working



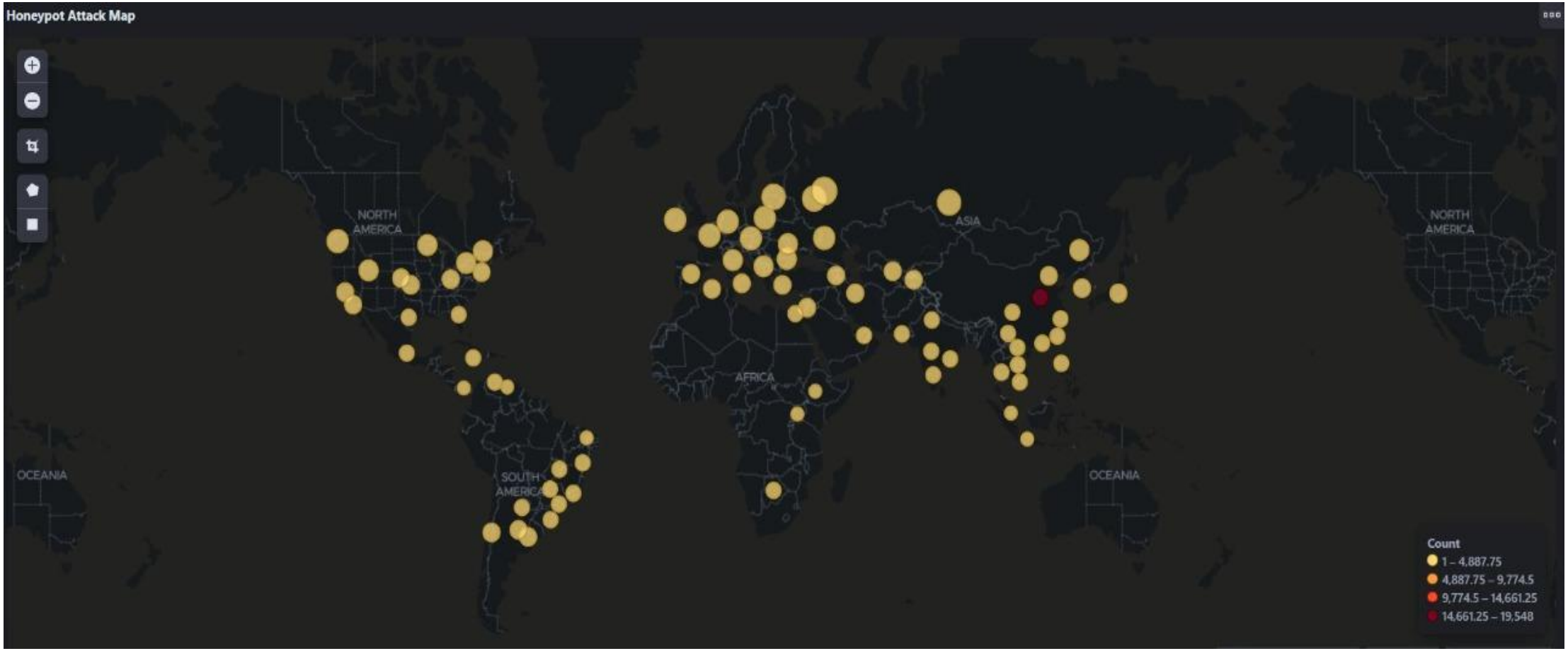
- Large volume of alerts will be incoming from two sources that are honeypot and internet.
- This alerts will be processed using Logstash, Metric Beat, File Beat, etc like components which are part of ELK stack.
- The more Threat Intelligence about the threat will be gathered from Threat Intelligence Repository.
- The Threat Intelligence Repository is powered by Spiderfoot, STIX Engine and Multilingual Threat Intelligence Translation.
- SIEM system will also perform the Real Time Analysis of Threat and automation of the event management using Kibana.
- Decisions on how to respond will be taken by Neural Network.
- Based on the same reports will be generated and populated to the user.

Prototype Design Demonstration



Tpot Mult-Honeypot Dashboard

Honeypot Attack Map



Attacks Distribution

Plan of Paper Publication

Artificial Intelligence based Security Orchestration, Automation and Response System accepted in “2021 6th International Conference for Converegence in Technology”.