Project Proposal

Computer Science Engineering Department

Project Title: Federated Learning Based Cyber Hygiene Tool

Project Owner: Aadhaar Koul (2020a1r040 – CSE – A1)

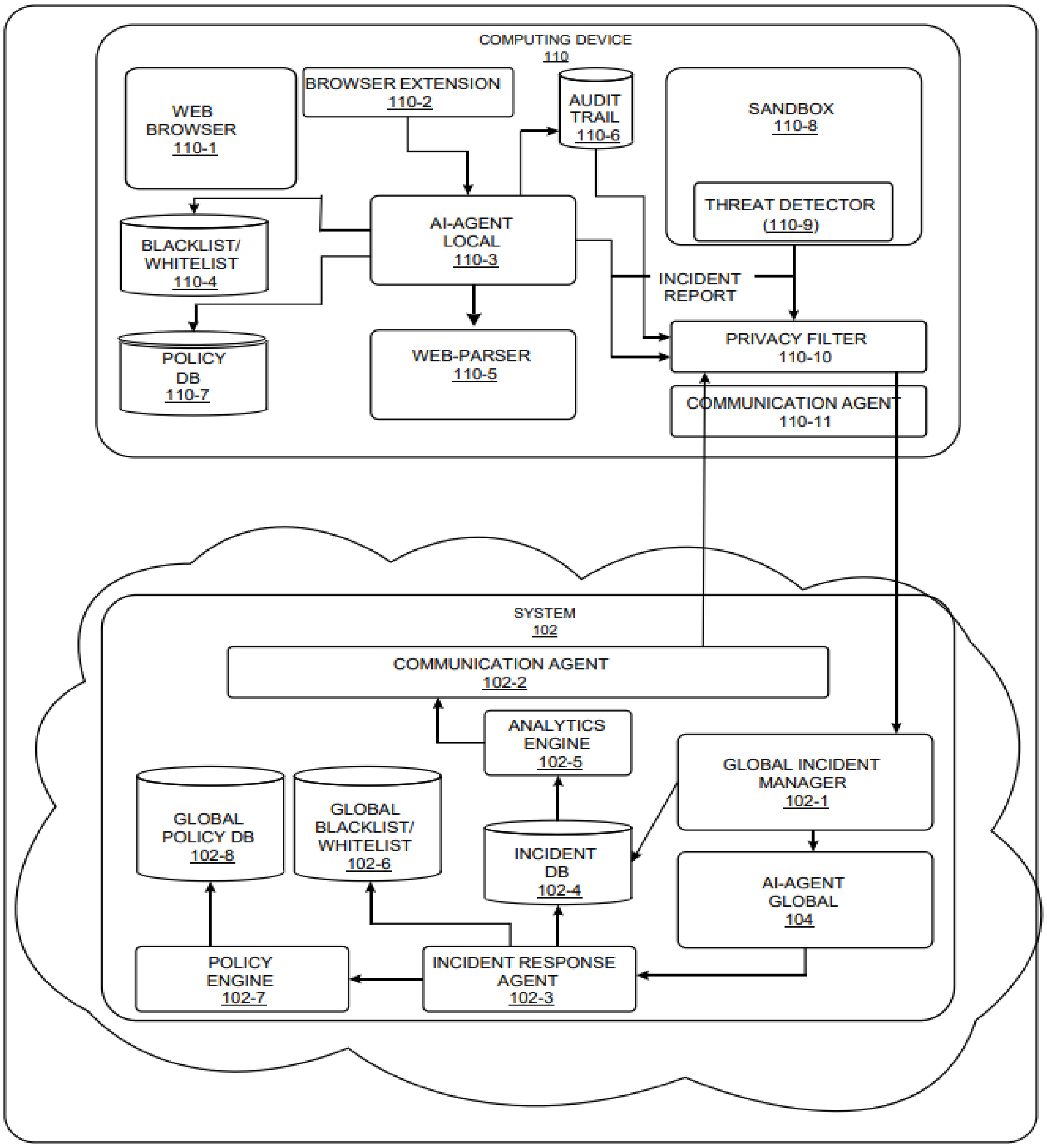
Abstract of Project:

Internet usage has become an integral part of our lives, and with it comes the risk of cyber threats. However, despite the growing awareness about cyber threats, many internet users remain unaware of the risks associated with their online activities. This naivety puts them at risk of falling victim to scams and frauds that can lead to significant financial losses.Recent studies have estimated that global cyber security threats may reach a staggering 10 trillion dollars. Unfortunately, there is no single solution that can fully protect internet users from these threats. As a result, many people remain susceptible to cyber-attacks, making it crucial to promote cyber hygiene awareness and educate people about safe internet practices.This abstract outlines the problem of the lack of cyber hygiene awareness among Indian internet users and the risks it poses. It highlights the need to promote awareness and implement protocols to safeguard users against cyber-attacks. Through research and analysis, we hope to identify the root causes of this problem and propose solutions that will help users protect themselves online. Our goal is to create a safer online environment for global internet users and reduce the risks associated with their online activities.

To address the problem of the lack of cyber hygiene awareness among Indian internet users and the increasing risks associated with online activities, we have developed a comprehensive solution called KAVACH. KAVACH is an AI-based advisory tool that provides users with a one-stop solution to safeguard their online activities. It keeps track of all online activities and provides real-time alerts and suggestions to users, enabling them to make informed decisions about their online behavior.KAVACH uses federated learning models to improve its algorithms continuously. By utilizing Random Forest, it identifies complex types of attacks and classifies them for the users, ensuring a safer online experience.Moreover, to ensure users' privacy and security, KAVACH uses a sandboxing agent that tests the product before deploying it onto users' devices. This approach enables us to detect and prevent malware attacks and prevent

cybercriminals from accessing users' confidential information.KAVACH is designed to meet the needs of users of all skill levels, making it easy for anyone to use it. Our ultimate goal is to create a safe online environment for Indian internet users and prevent them from falling victim to cyber-attacks. With KAVACH, we believe we can make a significant impact in promoting cyber hygiene awareness among Indian internet users and keep them safe online.[1]

Proposed Architecture:



Platform/Language Used: The Platform approved for this year's project includes the following programming languages - HTML , CSS , JS , Jquery , Python

Technology Domain: The proposed solution is aligned to Artificial Intelligence , Systems Security/Data or Information Security and Text Processing.

Comparative analysis of Existing Systems:

1. Federated-Learning-Based Anomaly Detection for IoT Security Attacks (done by Viraaji Mothukuri; Prachi Khare; Reza M. Parizi; Seyedamin Pouriyeh;)

In their research, they delve into the vulnerabilities of the Internet of Things (IoT), highlighting how the immense automation of tasks through IoT networks also poses a significant risk due to the vast collection of user data, which becomes an attractive target for malicious attacks. Traditional machine learning (ML) approaches in IoT security assume widespread availability and transferability of large training data, which conflicts with privacy concerns associated with user data.

To tackle this challenge, they propose a novel approach using federated learning (FL) for anomaly detection in IoT networks. Their method employs decentralized on-device data, utilizing federated training rounds on gated recurrent units (GRUs) models. This approach maintains data integrity on local IoT devices by sharing only learned weights with the FL central server. Additionally, an ensembler aggregates updates from various sources to enhance the accuracy of the global ML model.

Their experimental results demonstrate that this FL-based anomaly detection approach outperforms traditional centralized ML approaches in preserving user data privacy while achieving an optimal accuracy rate in detecting attacks within IoT networks.[2]

2. A Hybrid Approach to Privacy-Preserving Federated Learning

The research on federated learning, conducted by Stacey Truex, Nathalie Baracaldo, Ali Anwar, Thomas Steinke, Heiko Ludwig, Rui Zhang, and Yi Zhou, addresses the challenges of maintaining privacy in collaborative model training without sharing raw data. The study highlights that simply keeping data local during training isn't enough to ensure privacy, leading to recent attacks.

Existing federated learning methods using secure multiparty computation (SMC) are susceptible to inference, while differential privacy approaches may compromise accuracy, especially when dealing with numerous parties having relatively small data sets.

To tackle these issues, the paper proposes a novel approach combining differential privacy and SMC. This combination aims to minimize noise injection growth as the number of parties increases without compromising privacy, maintaining a specified trust level. This approach creates a scalable system that safeguards against inference threats while producing accurate models.

Moreover, the researchers validated their system's effectiveness by applying it to three different machine learning algorithms, demonstrating superior performance compared to current state-of-the-art solutions.[3]

Basic & Innovative Features in Project:

The extension has the following properties or features.

- The extension uses the blacklist filtering to remove the bad URLs.
- The extension uses alerting systems on certain types of sites that m ay have potential to be threatful in any possible nature.
- The extension allows the user to navigate to the authentic source of a software like VLC etc.
- The extension allows the user to be safe from any locally hosted servers like a reverse shell , a local server
- The extension saves the user from attending any http URL .
- The URL allows the user to be safe from any dark web portal that ends with a .onion subdomain.
- The extension allows the trusted sites like google etc. to pass through while maintaining the black and a white list simultaneously.
- The extension allows the user to keep a check on all the requested URLs in a GUI form that are often missed by he user .
- The extension saves the important information like the URL type , request id , get , code etc. in the gui form and in the chrome local storage.
- The user data does not leave his or her system rather the extension prepares a machine learning model that sends out the weights and behaviors that is accumulated while parsing through numerous URLs.

High Level Project Plan: Major phases of the project along with timelines and tasks assigned.

16th December - Local and global AI Integration
23rd December - Structured and seamless flow of data between extension and AI models.
26th December - Properly designed responsive GUIs.
29th December - Finetuning the project
30th December - Project Final Touches

References:

[1] V. Mothukuri, P. Khare, R. M. Parizi, S. Pouriyeh, A. Dehghantanha and G. Srivastava, "Federated-Learning-Based Anomaly Detection for IoT Security Attacks," in IEEE Internet of Things Journal, vol. 9, no. 4, pp. 2545-2554, 15 Feb.15, 2022, doi: 10.1109/JIOT.2021.3077803.

[2] Liang X, Zhang H, Tang W and Min F. (2024). Robust federated learning with voting and scaling. *Future Generation Computer Systems*. 10.1016/j.future.2023.11.015. **153**. (113-124). Online publication date: 1-Apr-2024.https://linkinghub.elsevier.com/retrieve/pii/S0167739X23004235

[3] AISec'19: Proceedings of the 12th ACM Workshop on Artificial Intelligence and SecurityNovember 2019Pages 1–11https://doi.org/10.1145/3338501.3357370