An Efficient Spam Detection Technique for IoT Devices using Machine Learning

ABSTRACT

The Internet of Things (IoT) is a group of millions of devices having sensors and actuators linked over wired or wireless channel for data transmission. IoT has grown rapidly over the past decade with more than 25 billion devices are expected to be connected by 2020. The volume of data released from these devices will increase many-fold in the years to come. In addition to an increased volume, the IoT devices produces a large amount of data with a number of different modalities having varying data quality defined by its speed in terms of time and position dependency. In such an environment, machine learning algorithms can play an important role in ensuring security and authorization based on biotechnology, anomalous detection to

improve the usability and security of IoT systems. On the other hand, attackers often view learning algorithms to exploit the vulnerabilities in smart IoT-based systems. Motivated from these, in this paper, we propose the security of the IoT devices by detecting spam using machine learning. To achieve this objective, Spam Detection in IoT using Machine Learning framework is proposed. In this framework, five machine learning models are evaluated using various metrics with a large collection of inputs features sets. Each model computes a spam score by considering the refined input features. This score depicts the trustworthiness of IoT device under various parameters. REFIT Smart Home dataset is used for the validation of proposed technique. The results obtained proves the effectiveness of the proposed scheme in comparison to the other existing schemes.

**EXISTING SYSTEM**

\_ Denial of service (DDoS) attacks: The attackers can flood the target database with unwanted requests to stop IoT devices from having access to various services. These malicious requests produced by a network of IoT devices are commonly known as bots [3]. DDoS can exhaust all the resources provided by the service provider. It can block authentic users and can make the network resource unavailable.

\_ RFID attacks: These are the attacks imposed at the physical layer of IoT device. This attack leads to loose the integrity of the device. Attackers attempt to modify the data either at the node storage or while it is in the transmission within network. The common attacks possible at the sensor node are attacks on availability, attacks on authenticity, attacks on confidentiality, Cryptography keys brute-forcing [4]. The countermeasures to ensure prevention of such attacks includes password protection,

data encryption and restricted access control.

\_ Internet attacks: The IoT device can stay connected with Internet to access various resources. The spammers who want to steal other systems information or want their

target website to be visited continuously, use spamming techniques [5]. The common technique used for the same is Ad fraud. It generates the artificial clicks at a targeted

website for monetary profit. Such practicing team is known as cyber criminals.

NFC attacks: These attacks are mainly concerned with electronic payment frauds. The possible attacks are unencrypted traffic, Eavesdropping, and Tag modification. The solution for this problem is the conditional privacy protection. So, the attacker fails to create the same profile with the help of user’s public key [6]. This model is based on random public keys by trusted service manager.

Disadvantages

* + In the existing work, the system is less effective due to lack of Spam Detection in IoT using Machine Learning framework.
  + This system is less performance in which it is clear that Supervised machine learning techniques is absence.

**PROPOSED SYSTEM**

The digital world is completely dependent upon the smart devices. The information retrieved from these devices should be spam free. The information retrieval from various IoT devices is a big challenge because it is collected from various domains. As there are multiple devices involved in IoT, so a large volume of data is generated having heterogeneity and variety. We can call this data as IoT data. IoT data has various features such as real-time, multi-source, rich and sparse..

**Advantages**

1) The proposed scheme of spam detection is validated using five different machine learning models.

2) An algorithm is proposed to compute the spamicity score of each model which is then used for detection and intelligent decision making.

3) Based upon the spamicity score computed in previous step, the reliability of IoT devices is analyzed using different evaluation metrics.

**SYSTEM REQUIREMENTS**

➢ **H/W System Configuration:-**

➢ Processor - Pentium –IV

➢ RAM - 4 GB (min)

➢ Hard Disk - 20 GB

➢ Key Board - Standard Windows Keyboard

➢ Mouse - Two or Three Button Mouse

➢ Monitor - SVGA

**SOFTWARE REQUIREMENTS:**

* **Operating system :** Windows 7 Ultimate.
* **Coding Language :** Python.
* **Front-End :** Python.
* **Back-End :** Django-ORM
* **Designing :** Html, css, javascript.
* **Data Base :** MySQL (WAMP Server).