



# A MACHINE LEARNING FRAMEWORK BASED ON VARIOUS NETWORK TRAFFIC CHARACTERISTICS TO IDENTIFY AND CLASSIFY THE DEFAULT BEHAVIOR OF IOT DEVICES ON A NETWORK

Pagalla Bhavani Shankar  
Department of CSE  
University College of Engineering & Technology  
Krishna University, Machilipatnam  
Andhra Pradesh, India

Yogi Reddy Maramreddy  
Department of CSE  
GITAM (Deemed to be) University  
Hyderabad  
Telangana, India

Padala S Venkata Durga Gayatri  
Department of CSE  
University College of Engineering  
Adikavi Nannaya University  
Rajamahendravaram, Andhra Pradesh, India

**Abstract---** The Internet of Things (IoT) is being well acquire to the next era of revolutionary generations amongst the new technologies. IoT technology being hailed so hard we had to stop in our society, smart homes, enterprises, and smart cities. Dynamics of smart one's are increasingly being equipped with a profusion of IoT devices. Due to the tremendous upgradation of knowledge in various aspects impresarios of such smart environments may not even be fully aware of their working nature or principles of IoT devices, assets and functioning properly safe from cyber-attacks. In this paper, we addressing this challenge by developing a robust framework for IoT device classification using traffic characteristics obtained at the level of network level. As a part of robust framework, firstly, we have a tendency to instrument a smart environment with 28 completely different IoT devices, spanning cameras, lights, plugs, motion sensors, appliances and health-monitors. We have a tendency to collect and synthesize traffic traces from this framework infrastructure for a period of 6 months, a type of subset of which we release as open data for the community to use. Second, we have to present or gifts the insights into the underlying network traffic characteristics using statistical and applied mathematical attributes such as activity cycles, port numbers, signaling patterns and cipher suites. Third, we have a tendency to develop a

**multi-stage machine learning based classification algorithm and demonstrate its ability to identify specific IoT devices with over 99% accuracy based on their network flow of activity. Finally, we have a tendency to discuss the trade-offs between cost, speed, and performance involved in deploying the classification network framework in real-time. Our study paves the way for impresarios of smart environments to monitor their IoT devices and assets for presence, functionality, and cyber-security without requiring any specialized devices or protocols.**

**Keywords – Internet of Things (IoT), Profusion, Impresarios, Cipher Suites, Cyber Security**

## I. INTRODUCTION

The number of devices connecting over the network of Internet is lightening at a peak, lead the way in the era of the "Internet of Things" (IoT). IoT refers to the greater number of low-cost devices that communicate and concatenate with each one in the network and with the remote servers on the Internet independently and autonomously. It comprises and includes daily needs of objects or things such as lights, cameras, motion sensors, door locks, thermostats, power switches and household appliances, with shipments projected to reach nearly 20 billion by 2020. Several number of IoT devices are expected to find their way in homes, enterprises,

campuses and cities, engendering “smart” environments benefiting the society and lives of people. Impresarios of smart environments can find it difficult to determine what IoT devices are connected to their network and further to ascertain whether each IoT device is functioning normally or not. This is the main attribute to the task of managing assets in an organization, which is typically distributed across different departments of the surroundings an IoT era.

## II. LITERATURE REVIEW

Internet of Things IoT is having wide range of applications in present trends [1,2]. Security Perspective is the key thing in Internet of Things [2]. Network or framework over the network leads to attainments of cyber security [3]. Ensuring of the security concerns will be Based on the group of remote servers and by the help of impresarios.

## III. INTERNET OF THINGS

Technology is a boon for us, to develop our selves and to develop our societies as per the needs of competitive world, by adopting the tremendous technologies. In this new era, a word is creating a magic and grab the world attentiveness from the all around the surroundings as in terms of IoT. It is a familiar word to present people and it seems to be easy to define as connecting, contacting and communicating between any objects over a network (Internet) is creates the principle of Internet of Things (IoT).

### OVER THE NETWORK OF INTERNET

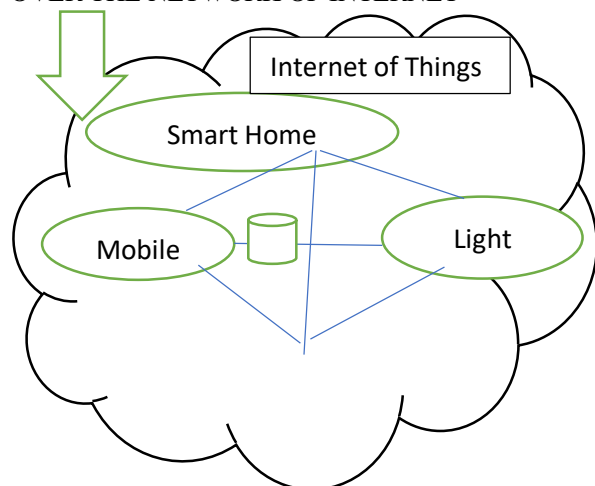


Fig. 1. Internet of Things (IoT) Environment

Fig. 1. delineate the meaningful diagrammatic representation of working principle of Internet of Things (IoT). Smart Home, Light, Mobile, Smart Cities are the things and the things are contacting, connecting and communicating with each other over the network of internet. By connecting, contacting

and communicating, things can collect the data and can share and exchange the data among the peers of things. Things are creating a type of network over the internet. By this rule of principle any one of the impresarios can detect and can trace the functionality of the things. The working principle of Internet of Things can change the working environment of impresarios. It is easier to detect the failure point of the thing in a network.

## IV. ARCHITECTURE OF IoT

Internet of Things (IoT) is a type of network-based protocol. IoT is mostly delineated with three-layer architecture: Perception Layer, Network Layer and Application Layer.

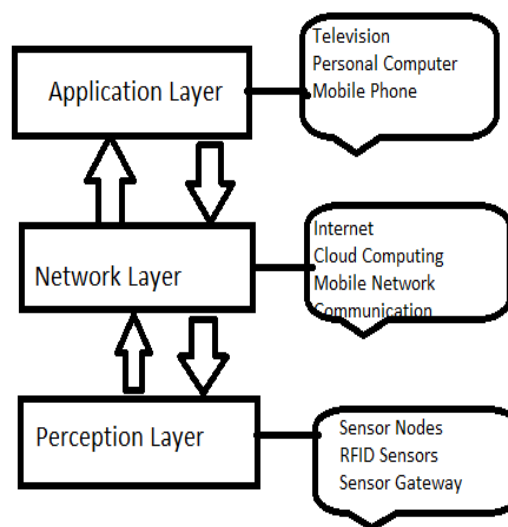


Fig. 2. Internet of Things (IoT) Layers Architecture

Perception Layer is the base layer of the architecture layers of Internet of Things. Perception Layer is also termed as Recognition Layer. The main scope of Recognition Layer is to display the properties of the objects and physical equipment of the objects within the network. Examples for the Perception Layer are Radio Frequency Identification (RFID) sensors, sensor gateways and sensor nodes.

Network Layer the layer, responsible for reliable transmission of the data or information over the network in the Internet of Things. Examples for reliable transmission over the network through internet, cloud computing and Mobile Network Communication.

Application Layer plays the top most priority among the all layers. Application Layer provides the personalized services to the users based on their choice by using an interface of Television, Personal Computer and Mobile Phone.

## V. APPLICATION OF IoT: EXAMPLE

Internet of Things (IoT) is an emerging technology and it is having several applications in all various domains likely, street lights, ground water, rainfall., etc.

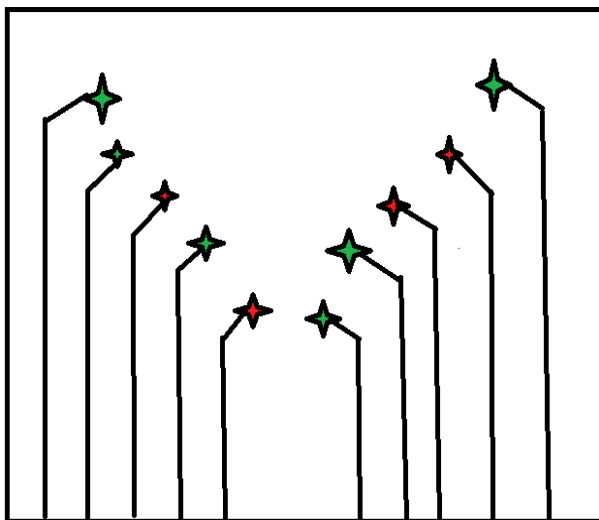


Fig. 3. Application of IoT: Monitoring Street Lights

Fig. 3. elucidate an example of IoT application of the monitoring the street lights. As an example, in Fig. 3. there are 10 street lights in a local area. Applying IoT to that local area, it is easier to monitor by the impresario. Reducing the manpower of impresario and effective faulting measure of a network of things will be evaluated efficiently. IoT will also give the economic feasibility. The Table. 1. will elucidate the status of the monitoring of the street lights.



Total No. of Street Lights	10
No. of Street Lights ON 	6
No. of Street Lights OFF / Defective 	2

Table. 1. Monitoring Status of Street Lights

### VI. DISCERN OF IoT

The escalation of IoT, however ensures an important problem. Impresarios of smart environments or smart cities will face a problem to find difficulty of the functioning of things in a proper manner or not. This is mainly attributed to the task of managing and organizing the assets in an organization, which is typically distributed across different departments of a network. Let an example, in a local council or body or organization, lighting sensors may be installed by the facilities team or impresarios, sewage and garbage sensors by the sanitation department and surveillance cameras by the local police division of that body. Coordinating across various departments to obtain an inventory of IoT assets is less time

consuming, onerous and error-prone, making it a like nearly impossible to know precisely what IoT devices are operating on the network at any point in time. Obtaining “visibility” into IoT devices or things in a timely predefined manner is of paramount importance to the impresarios, who is tasked with ensuring that things are in appropriate network security segments, were provisioned for requisite quality of service, and can be quarantined rapidly when breached at that level. The important credibility of visibility is emphasized in Cisco’s the most recent IoT security concern reports of documents and other, further highlighted by two recent events: sensors of a fish tank that compromised a casino in the month of July 2017 and attacks on a university campus network from its own vending machines in the month of February 2017. In these both cases, network segmentation could have potentially prevented or escaped from the attack and better visibility would have allowed rapid quarantining to limit the spoil or damage of the cyber-attack on the total overall enterprise network.

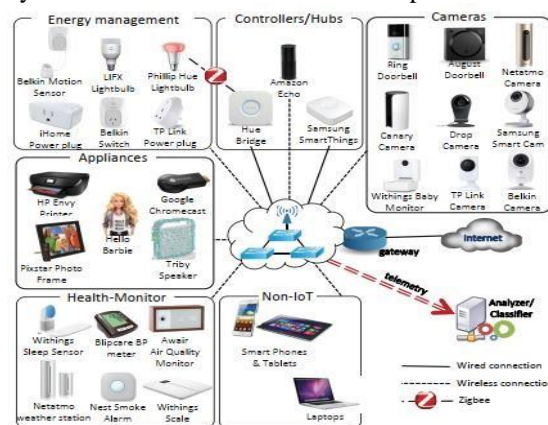


Fig. 4. Architecture of connected 28 different of IoT devices or things along with several non-IoT devices, and telemetry collected across the network of infrastructure is fed to our classification models

One would expect that devices can be recognized by their Media Access Control (MAC) address and Dynamic Host Configuration Protocol (DHCP) negotiation. However, it faces several challenges or issues: (a) IoT things manufacturers were typically uses NIC’s supplied or delivered by the third-party vendors, and hence the Organizationally Unique Identifier (OUI) prefix of the MAC address may not convey any information about the IoT device; (b) MAC addresses can be spoofed by malicious devices and by software; (c) many of the IoT things do not set the Host Name option in their DHCP requests indeed we found that about half the IoT devices we studied do not reveal their concerned specific host names.

Fig. 4. delineates connect the 28 different IOT devices and non-IOT devices its connected by a



network and its data sharing from devices to network through over the internet network.

## VII. IN LIGHT OF THE APPROACH

**Analysis of Empirical Traces:** The working nature of principle is one of the most first large-scale studies of reports to delve into the nature and characteristics of M2M traffic. It is inspired by the need to utilise understand whether M2M traffic imposes new challenges for the design model and management of cellular networks. The work uses a traffic trace of findings spanning one week from a tier-1 cellular network impresarios and compares M2M traffic with traditional smartphone traffic from a number of different perspectives of the nature – temporal variations, mobility, network performance, and so on. The reports of the study convey to shows the network impresarios to be cognizant of these factors when managing their networks models. In, the authors note that the amount of traffic generated by a single M2M device is likely to be small, but the total traffic generated by hundreds or thousands of M2M devices would be substantial. These observations from the report of studies are to some more extent corroborated by which note that a remote patient monitoring applications are expected to be to generate about 0.35 MB per day and smart meters roughly 0.07 MB per day in and around of the specified time intervals.

**Aggregated Traffic Model:** Aggregated Traffic Model is a type of coupled Markov Modulated Poisson Processes framework and it has to capture the behaviour of a single machine-type of network communication as well as the collective behaviour of the number of M2M devices or things in proposed model. The complexity of the XMPP framework is shown to grow linearly with the number of M2M devices, rendering it effective for large-scale synthesis of M2M traffic. A description of a model fitting is visualized via a use-case in fleet management comprising thousand trucks maintained by a transportation company or organization. And the fitting is purely based on measured M2M traffic from a 2G and 3G network. A simple and sample model to estimate the volume of M2M traffic generated in a wireless sensor network enabled connected home is constructed in and around the network. Yet, since behaviour of sensors are very application specific oriented, the work identifies certain common communication patterns that can be attributed to any sensor device. Using these kinds of attributes, four generalized substituted equations are proposed to estimate the volume of traffic generated by a sensor network enabled connected apartment or home or body. Use of Machine Learning strategy: Various machine-learning based analytical methods have been proposed in the literature to classify traffic

application or identify malwares/botnets for typical computer networks. The work in uses deep learning algorithms to classify flow types such as Hyper Text Transmission Protocol, Simple Message Transmission Protocol, Telnet, QUIC, Office365, YouTube and many more by considering six features namely source/destination port number, payload volume, Transmission Control Protocol (TCP) window size, inter-arrival durational time and directions of traffics that were extracted from the first twenty (20) packets from the overall 28 packets of a flow in a specified framework network. The overall work is carried out it is in suggests that botnets were exhibit identifiable as in traffic patterns that can be classified and ratified by considering features of the networks of packets such as average time between successive flows, flow duration, inbound and outbound traffic volume, and Fourier transformation (also be in Image Processing strategy) over the flow start times. And then the detection of malicious activity or source on the network was enhanced in by combining these flow level features with packet-level attributes including the packet size, the countered byte distribution of specified payload to a network of packets, inter arrival time duration of the packets and TLS handshake metadata (i.e., cipher suite codes). Further to this context of contents, authors have released an open source libpcap-based tool called Joy to extract these features from the passive capture of network traffic in a specified network of a framework. In the context and contents of IoT, it may use machine learning algorithms too to classify a single TCP flow control from authorized devices and things on the network frameworks. And it has employs over 300 functional attributes (both types of packet-level and flow-level), though the most and more influential one's are minimum or median or and also average of packets Time-To-Live (TTL), the ratio of total number of bytes transmitted and received in the network, overall total number packets with reset (RST) flag, and the Alexa rank of server. Enriching to this, while overall the above-mentioned semi-oriented works were making important and key contributions to the framework, they do not undertake as fine-grained characterization and classification of IoT things in a smart environment such as a home, city, campus, body, organization or enterprise. at a glance of addition to the working model of a framework, furthermore, statistical models were not developed and that enable IoT things classification based on their network traffic characteristics. Most importantly, prior of the works do not make any data set publicly available for the research community to use and build upon in a network. Our work will overcome these shortcomings.

## VIII. METHODOLOGY



New methodology has to be applied to the pre

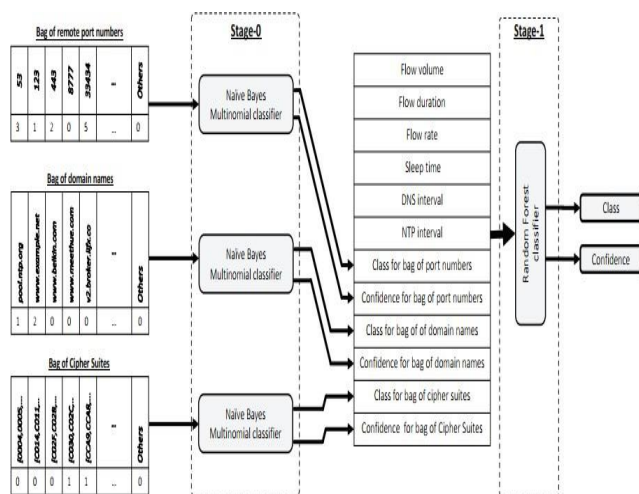


Fig. 5. Methodology approaches the network

working based model of the network architecture. By applying new methodology, there is a chance to be confine the functionality of the three attributes: set of domain names, set of remote port numbers and set of cipher suites in and as in a representation of matrix format. And also, there is a chance to refine the refining methods: server port numbers, DNS Queries, NTP Queries and Cipher suites.

## IX. CONCLUSION

Applying the new methodology to the pre working principle of a network, it yields to the better enhancement in finding the appropriate results and it enhances the working spot of cyber spots and enables the cipher suites at need of the network layers in a specified framework. It has to the fruitful approach to the future needs as per the user requirements.

## X. REFERENCES

- [1] Sunil Kumar Malge, Pallavi Singh, "Internet of Things IoT: Security Perspective".
- [2] Sachin Kumar, Pragma Tiwari, Mikhail Zymbler, "Internet of Things is a revolutionary approach for future technology enhancement: a review".
- [3] Pagalla Bhavani Shankar, "NSCS03: Enhancing approach to objective cyber security through digital literacy".
- [4] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley, "Color transfer between images," IEEE Computer Graphics Applications, vol. 21, no. 5, pp. 34–41, 2001.
- [5] S. Beigpour and J. van de Weijer, "Object recoloring based on intrinsic image estimation," ICCV, 2010.

- [6] F. Pitie, A. C. Kokaram, and R. Dahyot, "Automated colour grading using colour distribution transfer," Comput Vis Image Underst, pp. 123–137, 2007.
- [7] H. Chang, O. Fried, Y. Liu, S. DiVerdi, and A. Finkelstein, "Palette- based photo recoloring," ACM Transactions on Graphics (Proc. SIGGRAPH), vol. 34, no. 4, 2015.
- [8] M. P. Rao, A. N. Rajagopalan, and G. Seetharaman, "Harnessing motion blur to unveil splicing," IEEE Transactions on Information Forensics and Security, vol. 9, no. 4, pp. 583–595, 2014.
- [9] G. Muhammad, M. Hussain, and G. Bebis, "Passive copy move image forgery detection using undecimated dyadic wavelet transform," Digital Investigation, vol. 9, no. 1, pp. 49–57, 2012.
- [10] G. Cao, Y. Zhao, R. Ni, and X. Li, "Contrast enhancement-based forensicsindigitalimages," IEEE Transactions on Information Forensics and Security, vol. 9, no. 3, pp. 515–525, 2014.
- [11] X. Pan and S. Lyu, "Region duplication detection using image feature matching," IEEE Transactions on Information Forensics and Security, vol. 5, no. 4, pp. 857–867, 2010.
- [12] X. Zhao, J. Li, S. Li, and S. Wang, "Detecting digital image splicing in chroma spaces," in Digital Watermarking - International Workshop, 2010, pp. 12–22.
- [13] M. C. Stamm and K. J. R. Liu, "Forensic detection of image manipulation using statistical intrinsic fingerprints," IEEE Transactions on Information Forensics and Security, vol. 5, no. 3, pp. 492–506, 2010.
- [14] T. J. de Carvalho, F. A. Faria, H. Pedrini, R. da S. Torres, and A. Rocha, "Illuminant-based transformed spaces for image forensics," IEEE Transactions Inf. Forensics Security, vol. 11, no. 4, 2016.
- [15] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," IEEE Transactions Image Process, 2007.
- [16] J. S. Ho, O. C. Au, J. Zhou, and Y. Guo, "Inter-channel demosaicking traces for digital image forensics," in IEEE International Conference on Multimedia and Expo, 2010, pp. 1475–1480.
- [17] F. Pitie and A. C. Kokaram, "The linear monge-kantorovitch linear colour mapping for example-based colour transfer," IETCVMP, pp. 1–9, 2007.