Ransomware Auto-Detection In IoT Devices Using Machine Learning

Anshuman Dash
Reva University

Satyajit Pal
Reva University

Chinmay Hegde
Reva University

Ransomware Auto-Detection In IoT Devices Using Machine Learning

Anshuman Dash
REVA Academy for Corporate Excellence
REVA University
Bengaluru, India
anshumand.ba03@reva.edu.in

Satyajit Pal
REVA Academy for Corporate Excellence
REVA University
Bengaluru, India
satyajitp.ba03@reva.edu.in

Chinmay Hegde
REVA Academy for Corporate Excellence
REVA University
Bengaluru, India
chegde@astrikosconsulting.com

Abstract—**The term Internet of Things (often abbreviated IoT) was coined by industry researchers but has emerged into mainstream public view only more recently. The IoT is a massive group of devices containing sensors or actuators connected over wired or wireless networks. IoT has been rapidly growing over the past decade and, during the growth, security has been identified as one of the weakest areas in IoT. There are over six billion estimated devices currently connected to the Internet and an estimate of over 25 billion connected by 2020. IoT and its applications propagate to majority of life's infrastructure ranging from health and food production to smart cities and urban management.**

While efficiency and prevalence of IoT are increasing, security issues remain a necessary concern for industries. Internet connected devices, including those deployed in an IoT architecture, are increasingly targeted by cybercriminals due to their pervasiveness and the ability to use the compromised devices to further attack the underlying architecture. In the case of ransomware, for example, devices that can store a reasonably amount of data are likely to be targeted. Thus, ensuring the security of IoT nodes against threats such as malware is a topic of ongoing interest. While malware detection and mitigation research are now new, ransomware detection and mitigation remain challenging. Ransomware is a relatively new malware type that attempts to encrypt a compromised device's data using a strong encryption algorithm. The victim will then have to pay the ransom (usually using bitcoins) to obtain the password or decryption key. Consequences include temporary or permanent loss of sensitive information, disruption of regular operations, direct/indirect financial losses.

In this paper, we present a machine learning based approach to detect ransomware of IoT devices. Specifically, our proposed approach outperforms K-Nearest Neighbors, Neural Networks, Support Vector Machine and Random Forest, in terms of accuracy rate, recall rate and precision rate.

Keywords— ***Ransomware detection, Internet of Things (IoT) security, Machine learning, Malware detection, DDoS***

I. INTRODUCTION

The concept of Internet of things(IoT) was introduced by the growth of the widely used global network known as the internet along with the deployment of ubiquitous computing and mobiles in smart objects which brings new opportunities for the creation of innovative solutions to various aspects of life [1]. The concept of Internet of things(IoT)creates a network of objects that can communicate, interact and cooperate to reach a common goal [2]. IoT devices can enhance our daily lives, as each device stops acting as a single device and become part of an entire full connected system. This provides us with the resulting data to be analyzed for better decision making, tracking our businesses and monitoring our properties while we are far away from them [3].

While efficiency and prevalence of IoT are increasing, security issues remain a necessary concern for industries. Internet-connected devices, including those deployed in an IoT architecture, are increasingly targeted by cybercriminals due to their pervasiveness and the ability to use the compromised devices to further attack the underlying architecture [4]. In the case of ransomware, for example, devices that can store a reasonably amount of data (e.g., Android and iOS devices) are likely to be targeted. Thus, ensuring the security of IoT nodes against threats such as malware is a topic of ongoing interest.

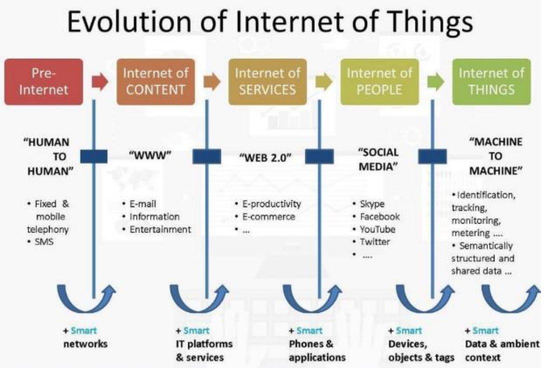


Fig. 1: Evolution of IoT
(https://twitter.com/fisher85m/status/926360908900773889)

II. IoT APPLICATIONS

As the paradigm of IoT is growing, it is stepping into every aspect of our lives. This leads to an easier life through wider range of applications such as smart city and its basic utilities.

A. Basic Utility Systems in Smart City

- *Smart Water Management System:* In this innovative day and age, the Internet of Things (IoT) is offering new solutions for improving water management to maximize efficient use. Comprehensive strategies utilizing the IoT can reduce water costs up to 20 percent. One of the reasons is how cheap IoT sensors are becoming with new battery-powered networking solutions (like LORA).

- *Smart Energy Management Systems:* Energy is a very important aspect for any household, industries, agriculture and so. Managing the energy efficiently and conserving it intelligently for appliances is very much important. The energy usage is directly affected with Coal, oil and so towards power generation.

- *Smart Gas Management Systems:* Urban natural gas provisioning has made significant progress in recent years. In the smart gas field, IoT enables stable, real-time traffic data collection from gas meters, device status monitoring, command delivery, and additional remote operations.

- *Smart Sewage Management Systems:* IOT in wastewater facility installs smart sensors at various points in their water management system. These sensors collect data on water quality, temperature variations, pressure changes, water, and chemical leaks, and they send those data back to a web application that synthesizes the information into actionable insights.

- *Smart Lighting Management Systems:* Improving safety within the city is one of the major benefits of smart lighting solution. The way smart lights work is that they derive power from the power grid and the street light poles have small cells that are a key enabler for smart solutions.

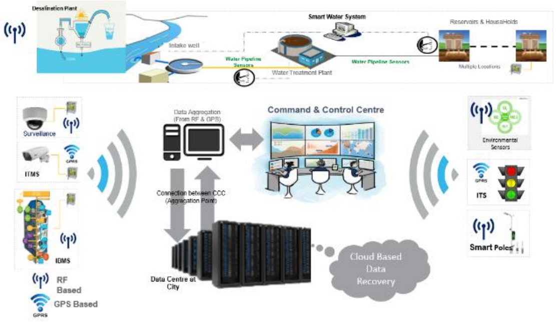


Fig. 2: An illustration of IoT based Smart City with Smart Utility Systems

B. *Safety & Security (Surveillance)*

A safer, mobilized version of our current cities where everything is automated through technologies such as artificial intelligence (AI), cloud, Big Data and analytics is what we envisage. But one of the key components that is almost forgotten is the importance of surveillance—24/7 round the clock monitoring of citizens, enabled with a robust infrastructure to ensure the safety and security of civilians.

- *Environmental sensors:* Sensors can measure environmental conditions such as air quality, temperature and humidity, water quality and noise levels, which are useful for city authorities to direct resources towards any mitigated imbalances.

- *Support to law enforcement:* Video surveillance feed has often been useful tool for law enforcement agencies to gather primary evidence during investigations. It is therefore critical for civic authorities to deploy high-quality CCTV cameras with required software interface to ensure that the recorded feed is of high-resolution.

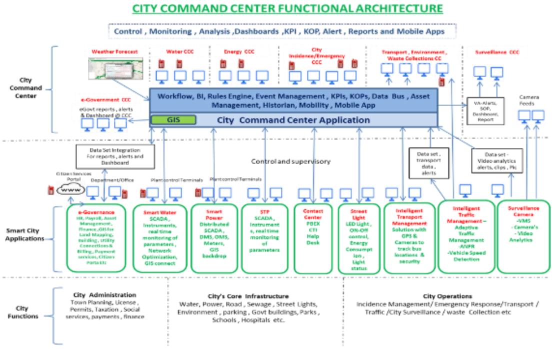

CITY COMMAND CENTER FUNCTIONAL ARCHITECTURE
Control , Monitoring , Analysis ,Dashboards ,KPI , KOP, Alert , Reports and Mobile Apps
City Command Center
Weather Forecast
Water CCC
Energy CCC
City Incidence/Emergency CCC
Transport , Environment , Waste Collections CC
Surveillance CCC
e-Government CCC
eGovt reports , alerts & Dashboard @ CCC
Workflow, BI, Rules Engine, Event Management , KPIs, KOPs, Data Bus , Asset Management, Historian, Mobility , Mobile App
GIS
City Command Center Application
VA-Alerts, SOP, Dashboard, Report
Camera Feeds
Data Set Integration For reports , alerts and Dashboard
Control and supervisory
Data set , transport data , alerts
Data set - Video analytics alerts, clips , Pic
Citizen Services Portal
www
Department/Office
Plant control Terminals
Plant control Terminals
Smart City Applications
e-Governance HR, Payroll, Asset Management, Finance ,GIS for Land Mapping, Building, Utility Connections & Billing, Payment services, Citizen Porta Etc
Smart Water SCADA , Instruments, real time monitoring of parameters , Network Optimization, GIS connect
Smart Power Distributed SCADA , DMS, OMS, Meters, GIS backdrop
STP SCADA , Instrument s, real time monitoring of parameters
Contact Center PBEX CTI Help Desk
Street Light LED Light , ON-Off control, Energy Consumpt ion , Light status
Intelligent Transport Management Solution with GPS & Cameras to track bus locations & security
Intelligent Traffic Management – Adaptive Traffic Management -ANPR -Vehicle Speed Detection
Surveillance Camera -VMS - Camera's - Video Analytics
City Functions
City Administration Town Planning, License , Permits, Taxation , Social services, payments , finance
City's Core Infrastructure Water, Power, Road , Sewage , Street Lights, Environment , parking , Govt buildings, Parks , Schools , Hospitals etc.
City Operations Incidence Management/ Emergency Response/Transport / Traffic /City Surveillance / waste Collection etc

Fig. 3: An illustration of IoT based Smart City Functional Architecture

(https://docs.google.com/presentation/d/1ck3o4M21qEvvXTLgXL56mdCwi6pvYlmITXNyv7JV0TU/edit#slide=id.g3e3c360a4a_0_27)

C. *Mobility*

- *Smart Traffic Management Systems:* Traffic management is one of the biggest infrastructure hurdles faced by developing countries today. Thus, the increased use of vehicles has caused an immense amount of traffic congestion. Several countries are overcoming this traffic bottleneck by fetching information from CCTV feeds and transmitting vehicle-related data to city traffic management centers to help create improvements.

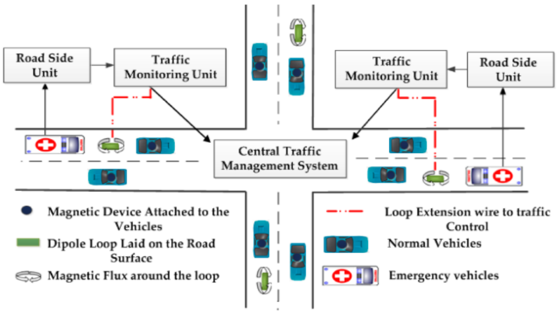

Road Side Unit
Traffic Monitoring Unit
Traffic Monitoring Unit
Road Side Unit
Central Traffic Management System
Magnetic Device Attached to the Vehicles
Dipole Loop Laid on the Road Surface
Magnetic Flux around the loop
Loop Extension wire to traffic Control
Normal Vehicles
Emergency vehicles

Fig. 4: An illustration of IoT based Smart Traffic Managements System

(https://www.mdpi.com/1424-8220/16/2/157)

- *Smart Transport Management Systems:* Smart transportation is developed on the base of smart infrastructure that includes not only multi-modal connected conveyance but also automated traffic signals, tolls and fare collection. Smart services offer different benefits, from smart parking and vehicle locating systems, to route diversion alerts.

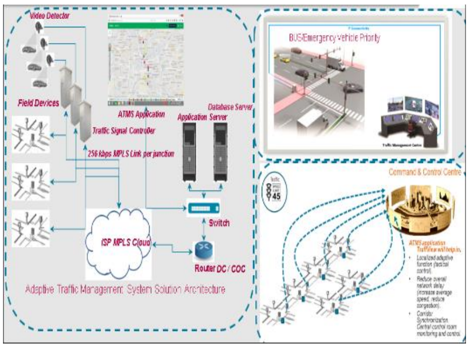

Video Detector
Field Devices
ATMS Application
Traffic Signal Controller
256 kbps MPLS Link per junction
Database Server
Application Server
Switch
ISP MPLS Cloud
Router DC / COC
Adaptive Traffic Management System Solution Architecture
BUS/Emergency Vehicle Priority
Traffic Management Centre
Command & Control Centre
45
ATMS application
Localized adaptive function (tactical control)
Reduce overall network delay (increase average speed, reduce congestion)
Corridor Synchronization Central control room monitoring and control

Fig. 5: An illustration of IoT based Smart Transport Management System

(https://docs.google.com/presentation/d/1ck3o4M21qEvvXTLgXL56mdCwi6pvYlmITXNyv7JV0TU/edit#slide=id.g3e3c360a4a_0_27)

D. *Governance*

Smart governance or good governance are two sides of the same coin. The use of the internet and digital technology is creating a progressive government- public partnership, strengthening government institutions and integrating all sections of society. To effectively manage these segments of society, our cities need smart administration and governance. Web portals, online forums, mobile apps and their unified services have helped public to directly share their questions, suggestions and grievances to government authorities. The forum on the website gives fellow applicants a space to share ideas, suggestions and know the status of other applications filed.

III. IoT Technologies And Protocols

Several Communication Protocols and Technology used in the internet of Things. Some of the major IoT technology and protocol (IoT Communication Protocols) are discussed here. These IoT communication protocols cater to and meet the specific functional requirement of an IoT system.

A. *6LOWPAN*

6LoWPAN is connecting more things to the cloud. Low-power, IP-driven nodes and large mesh network support make this technology a great option for Internet of Things (IoT) applications. As the full name implies – “IPv6 over Low-Power Wireless Personal Area Networks” – 6LoWPAN is a networking technology or adaptation layer that allows IPv6 packets to be carried efficiently within small link layer frames, such as those defined by IEEE 802.15.4 [7].

- *Network Architecture:* The uplink to the Internet is handled by the Access Point (AP) acting as an IPv6 router. Several different devices are connected to the AP in a typical setup, such as PCs, servers, etc. The 6LoWPAN network is connected to the IPv6 network using an edge router.

- *Security:* Security is a must for IoT systems and always presents a challenge. Due to the nature of IoT with many nodes that in many cases have very constrained performance, there are also more entry points for an outside attacker. Another critical aspect is that data flowing in a typical IoT system is not just “data,” the damage potential is much higher since the data flowing in the system can be used to open the door to you house or turn on/off alarms remotely.

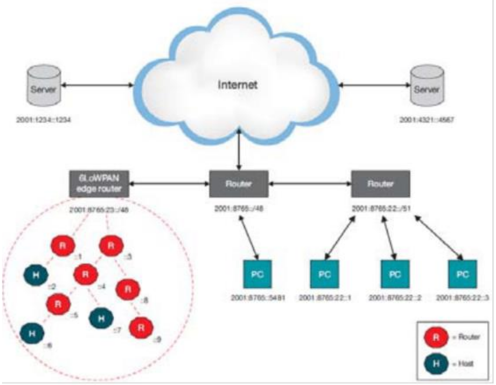


Fig. 6: 6LoWPAN Network Architecture

(https://www.researchgate.net/figure/An-example-of-an-IPv6-network-with-a-6LoWPAN-mesh-network-26_fig2_318075856)

B. *DASH7*

DASH7 Alliance protocol is an Actuator network protocol and it is an open source wireless sensor network. DASH7 operates in 433 MHz, 868 MHz and 915 MHz unlicensed ISM bands. DASH7 can provide a range up to 2km. For security AES 128-bit shared key encryption is used. Data rate offered by DASH7 is 28 kbps and it uses wakeup signal to achieve low power, low latency and elegant architecture.

DASH7 uses BLAST networking technology. BLAST means Bursty Light data Asynchronous and Transitive. In bursty, data transfer is abrupt and it does not contain contents like audio and video. In light, multiple consecutive packet transmission is generally avoided and the packet size is limited to 256 bytes. DASH7 communicates through command-response and therefore periodic hand shaking is not required.

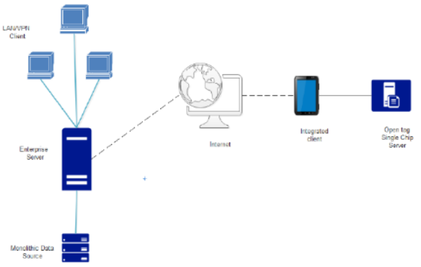


Fig. 7: DASH7 Network Architecture

(http://iopscience.iop.org/article/10.1088/1757-899X/396/1/012027/pdf)

C. *LoRa & LoRaWAN*

- *LoRa:* It is a radio modulation technology by Semtech Corporation. LoRa [3] provides connectivity about 15 to 20 km using chirp spread spectrum technique in which entire band width is used to transmit single signal. LoRa [8] uses 868 MHz to 900 MHz ISM bands for its operation and data rate is 0.3 kbps. Since LoRa have a long battery life, cost of replacement of devices can be reduced and its deployment is not so complex. LoRa use symmetric key cryptography to ensure security to its devices. Chirp spread spectrum used by LoRa can reduce the signal degradation, noise etc. while transmitting signal.

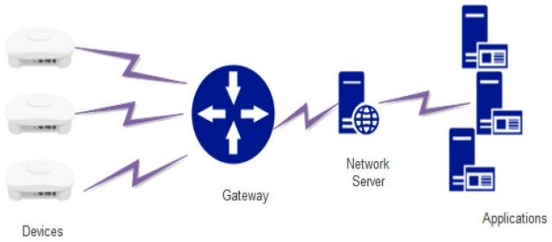


Fig. 8: LoRa Network Architecture

(http://iopscience.iop.org/article/10.1088/1757-899X/396/1/012027/pdf)

- *LoRaWAN:* LoRaWAN defines the communication protocol and system architecture for the network while the LoRa physical layer enables the long-range communication link. The protocol and network architecture have the most influence in determining the battery lifetime of a node, the network capacity, the quality of service, the security, and the variety of applications served by the network.

- *LoRaWAN Network Architecture:* Many existing deployed networks utilize a mesh network architecture. In a mesh network, the individual end-nodes forward the information of other nodes to increase the communication range and cell size of the network. While this increases the range, it also adds complexity, reduces network capacity, and reduces battery lifetime as nodes receive and forward information from other nodes that is likely irrelevant for them. Long range star architecture makes the most sense for preserving battery lifetime when long-range connectivity can be achieved.

- *Security:* LoRaWAN™ utilizes two layers of security: one for the network and one for the application. The network security ensures authenticity of the node in the network while the application layer of security ensures the network operator does not have access to the end user's application data. AES encryption is used with the key exchange utilizing an IEEE EUI64 identifier.

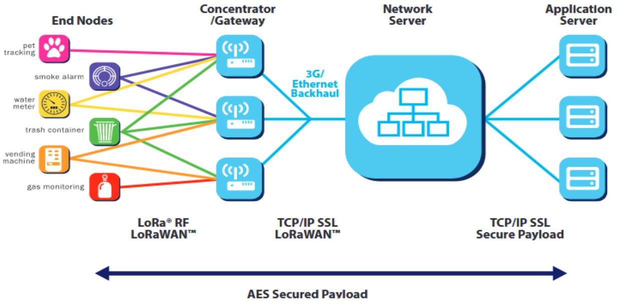


Fig. 9: LoRaWAN Network Architecture

(https://www.profissionaisti.com.br/2017/11/iot-protocolo-lorawan-e-principais-placas-de-desenvolvimento-lora/)

D. SigFox

Sigfox is a LPWAN technology which provides low power low data and low-cost communication devices. It operates in global network and is used for IoT communications. Even though there are plenty of wireless technologies Sigfox is widely accepted due its cheaper connection and an extended battery life. Sigfox enables new IoT applications and it can provide backup connectivity for higher bandwidth devices.

- *SigFox Network Architecture:* Communication using Sigfox can be performed when it detects an event or measure something, it will power on the communication module and send the message. The message is then picked up by the network and the data is received on your server. Sigfox is designed to maximize the energy efficiency. Sigfox consumes very low power when it transmits a data and no maintenance is required.
- *Security:* In Sigfox security challenge is addressed through a systematic process. Sigfox is one of the most secure LPWAN technology and it is unique in design. Sigfox devices predominantly operate in offline with a built-in behavior. When Sigfox needs to transmit or receive data from the internet, the Sigfox device will broadcast a radio message. This broadcasted message is received by base stations and the message is then transmitted to Sigfox core network, which is then delivered to corresponding IoT applications. This Sigfox network architecture provides an air gap and it is not possible to access an end through internet maliciously.

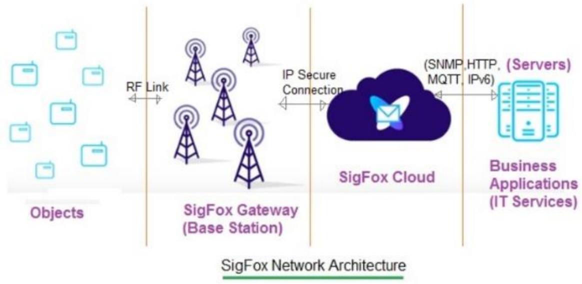

RF Link
IP Secure Connection
(SNMP, HTTP, MQTT, IPv6)
(Servers)
Objects
SigFox Gateway (Base Station)
SigFox Cloud
Business Applications (IT Services)
SigFox Network Architecture

Fig. 10: SigFox Network Architecture

(http://www.rfwireless-world.com/Tutorials/Sigfox-network-architecture.html)

E. Zigbee

Zigbee system structure consists of three different types of devices such as Zigbee coordinator, Router and End device. Every Zigbee network must consist of at least one coordinator which acts as a root and bridge of the network. The coordinator is responsible for handling and storing the information while performing receiving and transmitting data operations.

- *Zigbee Network Architecture:* Zigbee routers act as intermediary devices that permit data to pass to and from through them to other devices. End devices have limited functionality to communicate with the parent nodes such that the battery power is saved as shown in the figure. The number of routers, coordinators and end devices depends on the type of network such as star, tree and mesh networks.
- *Security:* The ZigBee standard includes complex security measures to ensure key establishment, secure networks, key transport and frame security. Those services are implemented at the Network and the Application Support Sublayer (APS), a sub layer of the Application Layer. The ZigBee protocol is based on an "open trust" model. This means all protocol stack layers trust each other. Therefore, cryptographic protection only occurs between devices. Every layer is responsible for the security of their respective frames.

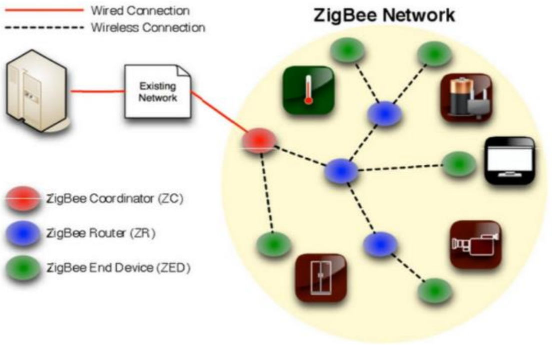


Fig. 11: Zigbee Network Architecture

(https://www.elprocus.com/what-is-zigbee-technology-architecture-and-its-applications/)

IV. Ransomware

Unlike traditional malware threats, a ransomware attack in IoT can be more devastating as it may affect an entire landscape of security services i.e., confidentiality, integrity, and availability, which may not only result in financial losses but may also result in an important information breach [5]. A ransomware may take entire control of data or a system and allow limited access for user interaction with the devices, ask for a hefty sum as a ransom, and release data to the user only after successful payment. In case a user does not pay, ransomware either extends the payment periods and ransom amount or deletes the data from the devices [6].

Initially, ransomware was named “AIDS”, as reported in 1989, when Joseph Popp distributed 20,000 infected floppy disk drives to the participants of World Health Organizations’ AIDS conference [8]. AIDS monitored the systems and counted number of times for which the systems were rebooted. It used to either encrypt data files or hide directory folders in C drive of infected computers. AIDS used to silently stay in the systems and get activated after a system reboots for 90 times.

A typical ransomware attack scenario involves infection of victim computer through penetration of an attack vector whereby the malware resulting from the attack contains a payload that, unbeknownst to the victim, engages in rendering important files as unusable, through their encryption with a key that is unknown to the victim [9]. Upon completion of the initial silent encryption phase, the original unencrypted files are deleted, and the victim is alerted that their files are now inaccessible and will remain so until a ransom is paid [10].

A. *Crypto Ransomware*

- Crypto-ransomware is a type of harmful program that encrypts files stored on a computer or mobile device to extort money. Encryption 'scrambles' the contents of a file, so that it is unreadable. To restore it for normal use, a decryption key is needed to 'unscramble' the file. Crypto-ransomware essentially takes the files hostage, demanding a ransom in exchange for the decryption key needed to restore the files.

- Unlike other threats, crypto-ransomware is neither subtle or hidden. Instead, it prominently displays lurid messages to call attention to itself, and explicitly uses shock and fear to pressure you into paying the ransom. A few so-called crypto-ransomware do not perform the encryption at all, and just use the threat of doing so to extort money. In most cases however, the threat is carried out.

B. *Locker Ransomware*

- Locker ransomware is a virus that infects PCs and locks the user’s files, preventing access to data and files located on the PC until a ransom or fines are paid. Locker demands a payment of $150 via Perfect Money or is a QIWI Visa Virtual Card number to unlock files. This variant affects Windows including Windows XP, Windows Vista, Windows 7, and Windows 8. Locker ransomware is a copycat of another very nasty ransomware that has infected over 250,000 computer systems named Crypto Locker. Although the Locker ransomware is simple, it can pack a devastating blow to one's computer.

C. *Hybrid Ransomware*

- Hybrid ransomware attacks that enable encryption and locking mechanisms are more dangerous because the device data and functionality could be compromised. A hybrid ransomware attack could become more vicious because it can target front-end and back-end IoT devices and systems.

V. MACHINE LEARNING ALGORITHMS

First, the overview of the machine learning field is discussed, followed by the description of methods relevant to this study. These methods include outperforms K-Nearest Neighbors, Neural Networks, Support Vector Machine and Random Forest, in terms of accuracy rate, recall rate and precision rate. The rapid development of data mining techniques and methods resulted in Machine Learning forming a separate field of Computer Science. The basic idea of any machine learning task is to train the model, based on some algorithm, to perform a certain task: classification, cauterization, regression, etc. Training is done based on the input dataset, and the model that is built is subsequently used to make predictions. The output of such model depends on the initial task and the implementation.

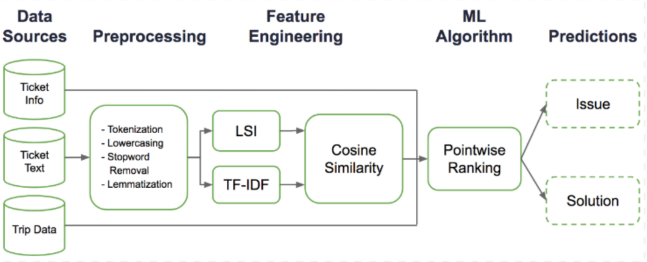


Fig. 12: General Workflow Of The Machine Learning Process

(https://eng.uber.com/cota/)

A. *Supervised and Unsupervised Learning*

- There are two machine learning approaches - supervised and unsupervised learning. In Supervised Learning, learning is based on labeled data. In this case, we have an initial dataset, where data samples are mapped to the correct outcome. The model is trained on this dataset, where it” knows” the correct results. In contrast to Supervised Learning, in Unsupervised Learning, there is no initial labeling of data. Here the goal is to find some pattern in the set of unsorted data, instead of predicting some value.

B. *Classification Methods*

- From machine learning perspective, Ransomware detection can be seen as a problem of classification or cauterization: unknown Ransomware types should be cauterized into several clusters, based on certain properties, identified by the algorithm. On the other hand, having trained a model on the wide dataset of malicious and benign files, we can reduce this problem to classification. For known Ransomware families, this problem can be narrowed down to classification only – having a limited set of classes, to one of which Ransomware sample certainly belongs, it is easier to identify the proper class, and the result would be more accurate than with cauterization algorithms.

- *K-Nearest Neighbors*: K-Nearest Neighbors (KNN) is one of the simplest, though, accurate machine learning algorithms. KNN is a non-parametric algorithm, meaning that it does not make any assumptions about the data structure. KNN can be used for both classification and regression problems. In both problems, the prediction is based on the k training instances that are closest to the input instance. In the KNN classification problem, the output would be a class, to which the input instance belongs, predicted by the majority vote of the k closest neighbors.

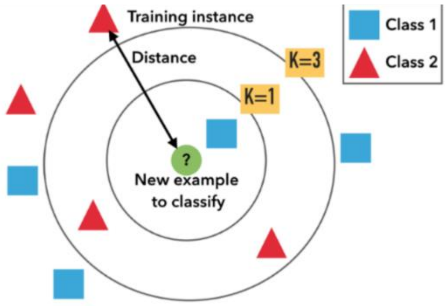


Fig. 13: KNN Example

(https://medium.com/@adi.bronshtein/a-quick-introduction-to-k-nearest-neighbors-algorithm-62214cea29c7)

$$Hamming\ Distance{:}\ d_{ij} = \sum_{k=1}^{p} |x_{ik} - x_{jk}|$$

$$Manhattan\ Distance{:}\ d_1(p,q) = ||p-q||_1 = \sum_{i=1}^{n} |p_i - q_i|$$

$$Minkowski\ Distance = \left(\sum_{i=1}^{n} |x_i - y_i|^p\right)^{1/p}$$

- The most used method for continuous variables is generally the Euclidean Distance, which is defined by the formulae below:

$$EuclidianDistance = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2}\ ;\ p\ and\ q\ are\ the\ points\ in\ n-space$$

- Euclidian distance is good for the problems, where the features are of the same type. For the features of different types, it is advised to use. The value of k plays a crucial role in the prediction accuracy of the algorithm. However, selecting the k value is a non-trivial task. Smaller values of k will most likely result in lower accuracy, especially in the datasets with much noise, since every instance of the training set now has a higher weight during the decision process. As a general approach, it is advised to select k using the formula below:

$$k = \sqrt{n}$$

- *Support Vector Machines:* Support Vector Machines (SVM) is another machine learning algorithm that is generally used for classification problems. The main idea relies on finding such a hyperplane, that would separate the classes in the best way. The term ’support vectors’ refers to the points lying closest to the hyperplane, that would change the hyperplane position if removed. The distance between the support vector and the hyperplane is referred to as margin.

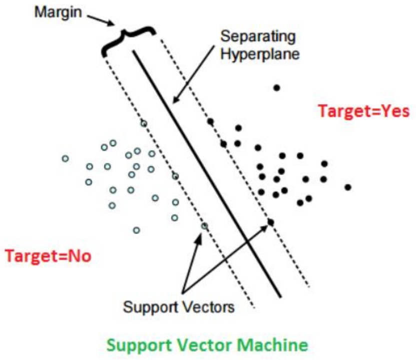


Fig. 14: SVM Example

(http://www.1proga.info/support-vector-machine-algorithm-389128593a10a109ec0dd4/)

- In the above figure, there is a dataset of two classes. Therefore, the problem lies in a two-dimensional space, and a hyperplane is represented as a line. In general, hyperplane can take as many dimensions as we want.
- The algorithm can be described as follows:
  - We define X and Y as the input and output sets respectively. $(x1,y1),\ldots,(xm,ym)$ is the training set.
  - Given x, we want to be able to predict y. We can refer to this problem as to learning the classifier y=f(x, a), where a is the parameter of the classification function.
  - F(x, a) can be learned by minimizing the training error of the function that learns on training data. Here, L is the loss function, and $Remp$ is referred to as empirical risk.

$$R_{emp}(a) = \frac{1}{m}\sum_{i=1}^{m} l(f(x_i,a),y_i) = Training\ Error$$

- We are aiming at minimizing the overall risk, too. Here, P(x,y) is the joint distribution function of x and y.

$$R(a) = \int l(f(x,a),y)dP(x,y) = Test\ Error$$

- We want to minimize the Training Error + Complexity term. So, we choose the set of hyperplanes, so f(x) = (w□x)+b:

$$\frac{1}{m}\sum_{i=1}^{m} l(w\cdot x_i + b, y_i) + ||w||^2\ \text{subject to}\ min_i|w\cdot x_i| = 1$$

- SVMs are generally able to result in good accuracy, especially on” clean” datasets. Moreover, it is good with working with the high-dimensional datasets, also when the number of dimensions is higher than the number of the samples. However, for large datasets with a lot of noise or overlapping classes, it can be more effective.
- *Random Forest:* Random Forest is one of the most popular machine learning algorithms. It requires almost no data preparation and modelling but usually results in accurate results. More specifically, Random

Forests are the collections of decision trees, producing a better prediction accuracy. That is why it is called a 'forest' – it is basically a set of decision trees. The basic idea is to grow multiple decision trees based on the independent subsets of the dataset. At each node, n variables out of the feature set are selected randomly, and the best split on these variables is found.

- Multiple trees are built roughly on the two third of the training data (62.3%). Data is chosen randomly.
- Several predictor variables are randomly selected out of all the predictor variables. Then, the best split on these selected variables is used to split the node. By default, the amount of the selected variables is the square root of the total number of all predictors for classification, and it is constant for all trees.
- Using the rest of the data, the misclassification rate is calculated. The total error rate is calculated as the overall out-of-bag error rate.
- Each trained tree gives its own classification result, giving its own" vote". The class that received the most" votes" is chosen as the result.

- *Neural Network:* Neural networks are one type of model for machine learning; they have been around for at least 50 years. The fundamental unit of a neural network is a node, which is loosely based on the biological neuron in the mammalian brain. The connections between neurons are also modelled on biological brains, as is the way these connections develop over time (with "training").

VI. USE CASE

A. *Discovery & Attack Surface Assesment*

- As connected IoT devices communicate, usually via HTTP initially, it's possible to continuously, automatically mine the involved user agent strings within the traffic, even if the devices communicate over nonstandard TCP ports, in order to discover these active devices and understand their device type, make, model, version, etc.

B. *Detection*

Continuously analysing the network traffic generated by IoT devices can help to identify potential compromise and misuse [11]. Here are some examples of traffic characteristics that can be monitored for:

- Repeated login attempts to IoT devices, which are using factory default credentials (e.g. admin/admin). This may indicate an attacker or malicious process is attempting to gain access via brute force.
- If IoT devices are moved to their own dedicated network segment / VLAN, monitor for unexpected traffic between the IoT devices and your other internal hosts on the main internal network.
- Outbound SOCKS proxy traffic from IoT devices.
- Communication from IoT devices to public IPs / Domains known to be associated with botnets, command and control, proxy usage, crypto mining, etc.
- The SSL/TLS certificates used to establish secure connections to public servers.
- Dynamically generated domains being looked up from IoT devices.
- Changes in traffic rates and patterns from an established baseline.

As noted earlier, IoT devices must communicate across the network to function and attackers must communicate across the network to repurpose them. With network traffic analysis, security analysts have the means to identify devices of interest and examine the full extent of their activity, even amongst the din of an internet of things.

C. *Metrics And Cross Validation*

We use the following four common performance indicators for malware detection:

- True positive (TP): indicates that a ransomware is correctly predicted as a malicious application.
- True negative (TN): indicates that a goodware is detected as a non-malicious application correctly.
- False positive (FP): indicates that a goodware is mistakenly detected as a malicious application.
- False negative (FN): indicates that a ransomware is not detected and labelled as a non-malicious application.

To evaluate the effectiveness of our proposed method, we used machine learning performance evaluation metrics that are commonly used in the literature, namely: Accuracy, Recall, Precision and AUC. *Accuracy* is the number of samples that a classifier correctly detects, divided by the number of all ransomware and goodware applications:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

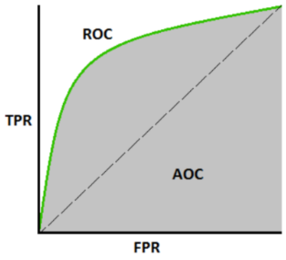
Recall or detection rate is the ratio of ransomware samples that are correctly predicted, and is defined as follows:

$$Recall = \frac{TP}{TP + FN}$$

Precision is the ratio of predicted ransomware that are correctly labelled a malware. Thus, Precision is defined as follows:

$$Precision = \frac{TP}{TP + FP}$$

AUC (Area Under the Curve) represents the probability that a true positive (green) example is positioned to the right of a true negative (red) example. AUC ranges in value from 0 to 1. A model whose predictions are 100% wrong has an AUC of 0.0; one whose predictions are 100% correct has an AUC of 1.0.



- *Performance of Algorithms*: We will now use the leave-one-out technique for cross validation. Below figure illustrate the network traffic usage graph due to ransomware.

D. Data Flow

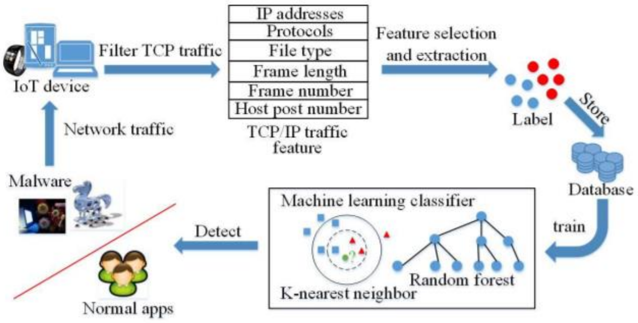
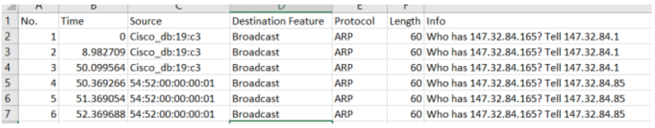


Fig. 14: Data Flow Diagram

(http://www.1proga.info/support-vector-machine-algorithm-389128593a10a109ec0dd4/)

The IoT device filters the TCP packets and selects the features among various network features including the frame number and length, labels them and stores these features in the database. The K-NN based ransomware detection assigns the network traffic to the class with the largest number of objects among its K nearest neighbours. The random forest classifier builds the decision trees with the labelled network traffic to distinguish ransomware. According to the experiments in [14], the true positive rate of the K-NN based ransomware detection and random forest-based scheme with dataset are 93% and 69%, respectively. In a ransomware detection scheme as developed in [15], an IoT device can apply the Q-learning to achieve the optimal offloading rate without knowing the trace generation and the radio bandwidth model of the neighbouring IoT devices.

- Network Data Set collected from the IOT devices in pcap file format. Below is a sample of the data which was converted and extracted using wireshark from the pcap file format.



	A	B	C	D	E	F	
1	No.	Time	Source	Destination Feature	Protocol	Length	Info
2	1	0	Cisco_db:19:c3	Broadcast	ARP	60	Who has 147.32.84.165? Tell 147.32.84.1
3	2	8.982709	Cisco_db:19:c3	Broadcast	ARP	60	Who has 147.32.84.165? Tell 147.32.84.1
4	3	50.099564	Cisco_db:19:c3	Broadcast	ARP	60	Who has 147.32.84.165? Tell 147.32.84.1
5	4	50.369266	54:52:00:00:00:01	Broadcast	ARP	60	Who has 147.32.84.165? Tell 147.32.84.85
6	5	51.369054	54:52:00:00:00:01	Broadcast	ARP	60	Who has 147.32.84.165? Tell 147.32.84.85
7	6	52.369688	54:52:00:00:00:01	Broadcast	ARP	60	Who has 147.32.84.165? Tell 147.32.84.85

Threats are clustered according to attack type: Denial of service (DoS), where some resource is swamped, causing DoS to legitimate users. Probes, gathering network information to bypass security. The data set is divided into two sets while executing the model as train and test. The division of the data is chosen in a way most demanding for classifiers. Train and test data do not show the same attack probability distribution; moreover, 16 out of the 38 labelled threats are only present in the test data set.

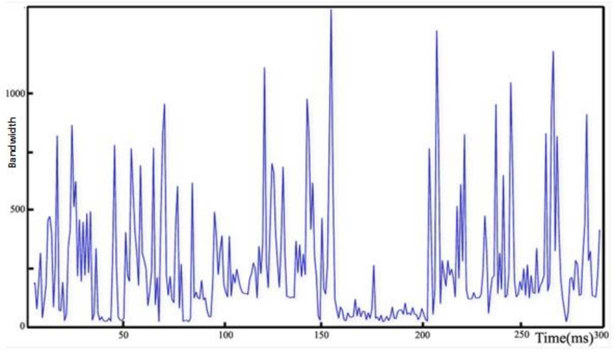


Fig-15 Benign Data

- Network Traffic data analysis on benign data given below. [Fig-15]

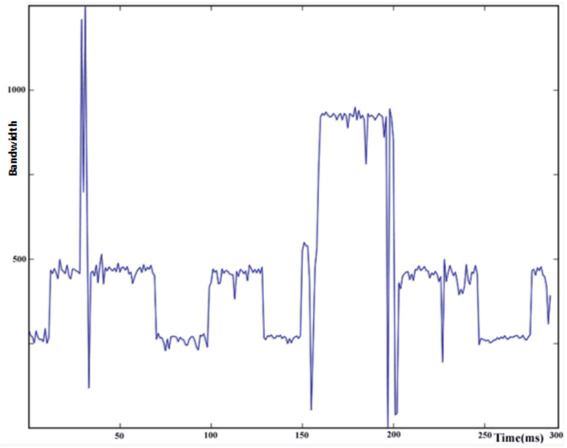
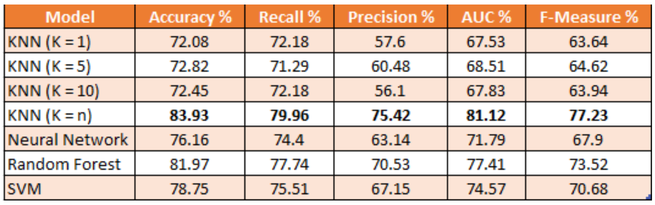


Fig-16 Infected Data with Ransomware

- Network Traffic data analysis on infected data with Ransomware given above. [Fig-16].

Comparison of the above two figures reflects a significant difference between patterns of network usage for ransomware versus benign applications.

- *Performance of Machine Learning Techniques: A Comparative Summary*



Model	Accuracy %	Recall %	Precision %	AUC %	F-Measure %
KNN (K = 1)	72.08	72.18	57.6	67.53	63.64
KNN (K = 5)	72.82	71.29	60.48	68.51	64.62
KNN (K = 10)	72.45	72.18	56.1	67.83	63.94
KNN (K = n)	**83.93**	**79.96**	**75.42**	**81.12**	**77.23**
Neural Network	76.16	74.4	63.14	71.79	67.9
Random Forest	81.97	77.74	70.53	77.41	73.52
SVM	78.75	75.51	67.15	74.57	70.68

- Evaluation Metrics for different Window Sizes, KNN and Euclidean distance: A Comparative Summary given below.

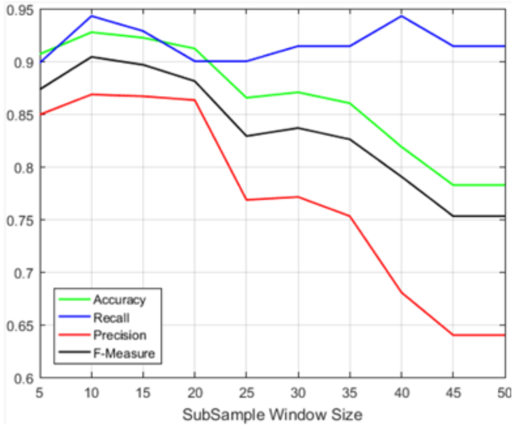


Fig-17

- Evaluation metrics for different window sizes, KNN and n distance: a comparative summary

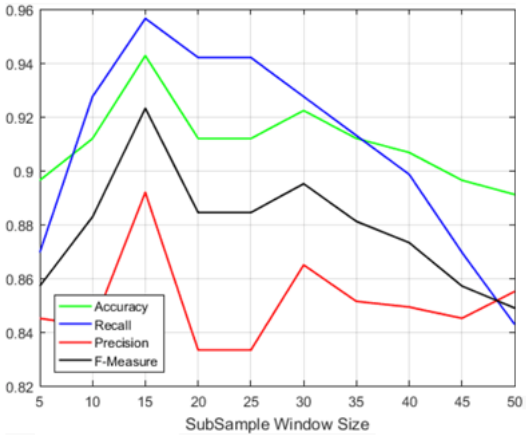


Fig-18

Furthermore and in practice, KNN's requirement for concurrent distance calculations between training and testing objects can be implemented using parallel processing (so distances can be independently computed). Subsamples dictionary can be partitioned into sperate IoT nodes and each subsample is sent to nodes. They return a label and a similarity value and the label having less similarity value is final subsample's label. This approach reduces the classification time and mitigates the need for storage capacity in every node.

VII. Future Directions of IoT Technologies In Smart City

IoT technologies contribute significantly to the majority of the detailed aspects of smart city technologies and infrastructure. Because the fundamental concepts and ideas of IoT technologies are shared with those of smart city technologies and infrastructure, a large number of business opportunities and extensive growth potential exist. Moreover, for the efficient and successful development of future IoT technologies in smart cities, the following points require focus.

Importance and essentiality: Sensor-oriented technologies for wireless networking are considered the top priority of IoT technologies for a smart city infrastructure.

- Importance: In addition to sensor-oriented technologies, technologies for network services are considered the most important IoT technologies for a smart city infrastructure.
- Essentiality: Compared to other technologies, energy-related technologies are considered the most essential IoT technologies for the smart city infrastructure. Because the technologies applied in a smart home environment are fundamental aspects of a smart city, the technologies and infrastructure for a smart home network should be swiftly developed and prepared.

In addition to these points, there are some challenging aspects that should also be resolved. First, because IoT technologies should provide various operating systems, which includes low to high-capacity processors, providing appropriately distributed resources is one of the most important tasks required in devices employing IoT technologies.

Second, data management solutions are required for the massive amounts of data collected by various IoT technology devices because the majority of such data is unstructured or atypical. This means that technologies for data categorization and intelligent analysis should be developed and introduced.

Third, the current IoT technology services are provided through independent specialized solutions, which are oriented and operated within a specific environment. Therefore, compatible integrated Sustainability applications for providing various IoT technology services should be developed and prepared by using appropriate network technologies.

Fourth, both appropriate solutions and plans for data security and privacy should be established. When users connect to an IoT technology service, reliable data processing and storage should be applied with confidentiality, integrity, and privacy. This means that reliable and safe communications and connections from each IoT technology device to the smart city infrastructure should be provided.

VIII. Conclusion

With increasing prevalence of Internet-connected devices and things in our data-centric society, ensuring the security of IoT networks is vital. Successfully compromised IoT nodes could hold the network to ransom. For example, in the case of ransomware, denying availability to data in an IoT network could adversely affect the operation of an organisation and result in significant financial loss and reputation damage.

In this paper, we presented an approach to detect ransomware, using network traffic flow analysis. Specifically, we utilize the unique local fingerprint of ransomware's network usage pattern to distinguish ransomware from non-malicious applications. Our set of experiments demonstrated that our approach achieved a detection rate of 93.76% and a precision rate of 89.85%.

We have shown that ML is a viable and effective approach to detect new variants and families of ransomware for subsequent analysis and signature extraction, and as a complement for AV. Mutual Information has shown to be an effective way of automatically selecting the features, while Regularized K Nearest Neighbors has shown to be an accurate algorithm, easy to train and update, and fast.

References

[1] [Online] Available: https://iot-analytics.com/internet-of-things-definition/.

[2] [Online] Available: http://www.gartner.com/newsroom/id/3165317.

[3] Tankard, Colin. "The security issues of the Internet of Things." Computer Fraud & Security 2015, no. 9 (2015): 11-14.

[4] Watson S, Dehghantanha A (2016) Digital forensics: the missing piece of the internet of things promise. Comput Fraud Secur 2016(6):5–8

[5] O'Grman G, McDonald G (2012) Ransomware: a growing menace. Tech. rep., Symantec Corporation. http://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/ransomware-a-growing-menace.pdf. Accessed 12 Feb 2017

[6] Caviglione L, Gaggero M, Lalande JF, Mazurczyk W, Urbański M (2016) Seeing the unseen: revealing mobile malware hidden communications via energy consumption and artificial intelligence. IEEE Trans Inf Forensics Secur 11(4):799–810

[7] Jonas Olsson, Texas Instrument, 6LoWPAN demystified, http://www.ti.com/lit/wp/swry013/swry013.pdf

[8] E. Bertino and N. Islam, "Botnets and internet of things security," Computer, vol. 50, pp. 76–79, 2017.[24] B. Nassi, A. Shamir, and Y. Elovici, "Oops!... i think I scanned a malware," arXiv preprint arXiv:1703.07751, 2017.

[9] R. Richardson and M. North, "Ransomware: Evolution, mitigation and prevention," International Management Re-view, vol. 13, no. 1, p. 10, 2017.

[10] The rise of ransomware and emerging security challenges in the Internet of Things. Available from: https://www.researchgate.net/publication/319527564_The_rise_of_ra

nsomware_and_emerging_security_challenges_in_the_Internet_of_Things [accessed Oct 28, 2018].

[11] [Online] https://www.corvil.com/blog/2018/discovering-iot-devices-and-monitoring-their-network-traffic

[12] Haykin S (1998) Neural networks: a comprehensive foundation, 2nd edn. Prentice Hall, 20(8):2481–2501 Kim H, Smith J, Shin KG (2008) Detecting energy-greedy anomalies and mobile malware variants. In: Proceedings of the 6th international conference on mobile systems, applications, and services. ACM, pp 239–252

[13] Kohavi R et al (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. Ijcai 14:1137–1145

and mobile malware variants. In: Proceedings of the 6th international conference on mobile systems, applications, and services.

[14] F. A. Narudin, A. Feizollah, N. B. Anuar, and A. Gani, “Evaluation of machine learning classifiers for mobile malware detection,” Soft Computing, vol. 20, no. 1, pp. 343–357, Jan. 2016.

[15] L. Xiao, Y. Li, X. Huang, and X. J. Du, “Cloud-based malware detection game for mobile devices with offloading,”

IEEE Trans. Mobile Computing, vol. 16, no. 10, pp. 2742–2750, Oct. 2017.