ISSN No:-2456-2165

Detection and Prevention of Advanced Persistent Threat (APT)

[1]Shambhavi I pattar, [2]Anand Pashupatimath
[1]M.Tech student, [2]Asst.Professor
[1, 2]Department of Computer Science and Engineering,
SDM College of Engineering and Technology, Dharwad

Abstract:- From some recent years maintaining security in computer system leads many problems by the APT attacks. These APT attacks are widely impacted in various industries. The team advanced persistent threats refers to well-formed, malicious group of persons who targets to attacks against specific computer systems. The attacks themselves enduring, difficult to find and most often use very advanced hacking techniques. There are many methods to increase the security level of organization in order to maintain these incidents. First approach is improving administrator's education about these kinds of security issues. Second is developing strict security policies. That includes access control and restrictions (to information or network), protecting information by encrypting it and installing latest security upgrades.

Keywords: - *RISC, MIPS, Xilinx Tool, Verilog-HDL, RTL, FPGA*

I. INTRODUCTION

In today's world one of the principle challenge in data security is new kind of threat that has turned into a huge uncertainty for different industries. They are adequate and contain assets to dispatch refined attacks. Hence, it got entitled as advanced persistent threats (APT). **Network:** In order to share the system resources, two or more computers are connected. The network may be linked through satellites, cables etc. There are varieties of network such as Wireless LAN, Metropolitan Area Networks, and Wireless WAN. **Threat:** A threat is anything that may possibly cause deliberate vandalism to the system. A threat leads to serious intrusion on systems. Some types of threats are: **Worms**: This virulent program escapade operating system amenability to spread itself. It is named as worms because the attack crawls like worm from one system to another. **Trojans**: Programs carrying unauthorized actions on computers, such as removing information from drives, taking out confidential information etc. Trojans aren't viruses. Rather than entering on own, they are merely spread by hackers. **Viruses:** These are programs infecting another program and summing own code to acquire control over files whenever they're opened. Advanced Persistent Threats is a well co-ordinate, group of people launching attacks on the required target. The threats imposed are long lasting leading to data loss and corruption of the system. Business operations are also damaged using APT. Sensitive information is also lost due to these kinds of attacks. Attack methodologies are changed by introducing APT. No system can be said as secure system, but the risks can be reduced. FireEye is an US based network security system that has detected more than 4000 attacks only related to APT. It is also found that APT has also caused malware infections. Basic prevention measures can be taken to avoid attacks such as educating user, protecting valuable data, block listing unknown malware servers etc. Even after knowing that there is no guarantee of any network being secured, companies can implement policies that can at least reduce risks of any external attack. Misuse of zero-day vulnerabilities in APT attacks in especially stressing amid 2013, FireEye detailed 11 attacks of that categories by java's zero-day

Vulnerabilities followed by Internet browsers and adobe reader exploited. Most of these category attacks are APT attacks caused damages to the industries reputation around $9.0 million amount.APT attacks well-organized attacks frequently stay undetected for a significant of time. APTs come with different attack methodology. This different attack methodology brings changes in information security of organizations.

II. ADVANCE PERSISTENT THREAT

This refers to the sensible and group of intruders who introduce attacks in computer systems such as government, IT companies or military. These attacks are enduring and use leading mutilate techniques. Since they are advanced, lengthened and constant in nature the organizations must have immense learning about the tools and execute them. These kinds of attacks are commonly implemented in some phases such as reconnaissance, preparation, execution, gaining access, information gathering and connection maintenance. In all of these steps invasion can be found with different probabilities. We need to teach the clients about various attacks and we should provide them sufficient knowledge so that the attacks can be reduced. Secondly, the authorities must accomplish rigid security approaches.

A. Phishing:

Phishing: Phishing is the criminal endeavor to get delicate subtleties, for example, passwords, Visa data and username. It's typically done by messaging or email spoofing; it regularly guides clients to enter individual data at a phony site which coordinates the look and feel of the legitimate site. Attempts to manage phishing incidents incorporate enactment, client training, and specialized safety efforts. Phishing attack can have many types, namely Whaling attacks, Pharming, Voice Phishing, and Spear

Phishing. In our project we proposed Spear Phishing attack, we will see spear phishing in detail in next part of our project paper.

➢ *Spear Phishing:*

These attacks are aimed particular persons or companies, regularly using information to the targeted that has been collected to more successfully repent the email as being authentic. Spear phishing online electronic mails might contain references to administrative or coworker at the victim's company, just as the utilization of the victim's individual name, are other individual information. Spear phishing attack is differ then other phishing attacks, and these attacks are more often attacks now days, so we choose spear phishing attack specifically in phishing attacks.

III. METHODOLOGY

A. *Architecture Design*

The figure 1 depicts the methodology for the project i.e. it provides information about various modules involved in the project.

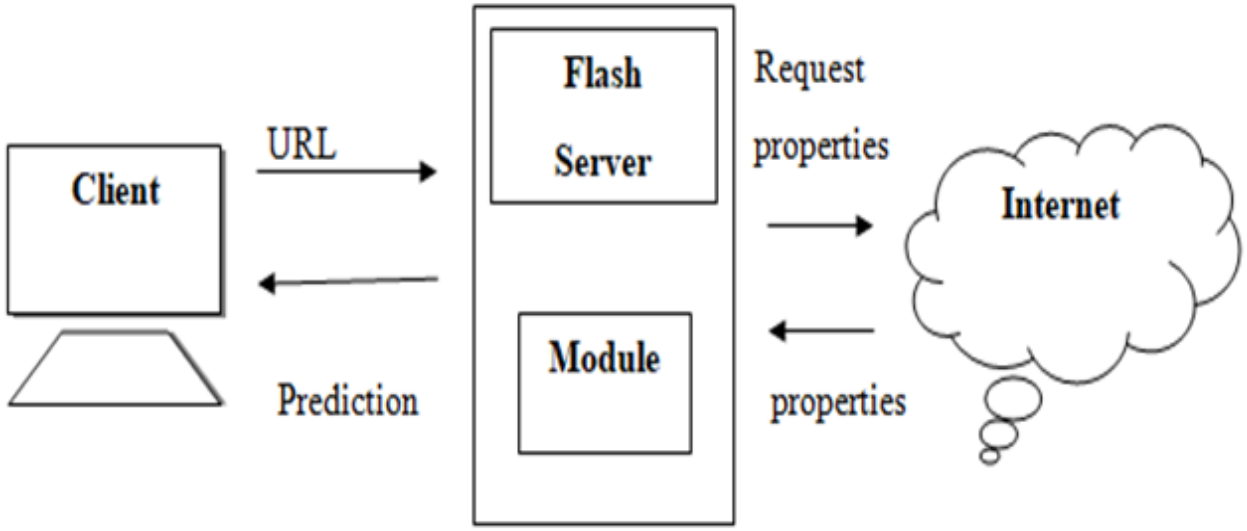

Fig 1:- Architecture Design

B. *Client:*

Whenever client receiver's some URL's through the mail, he adds URL to the flash server to know the URL is legitimate or not, because to protect the secured information about organization. For example consider figure 2.

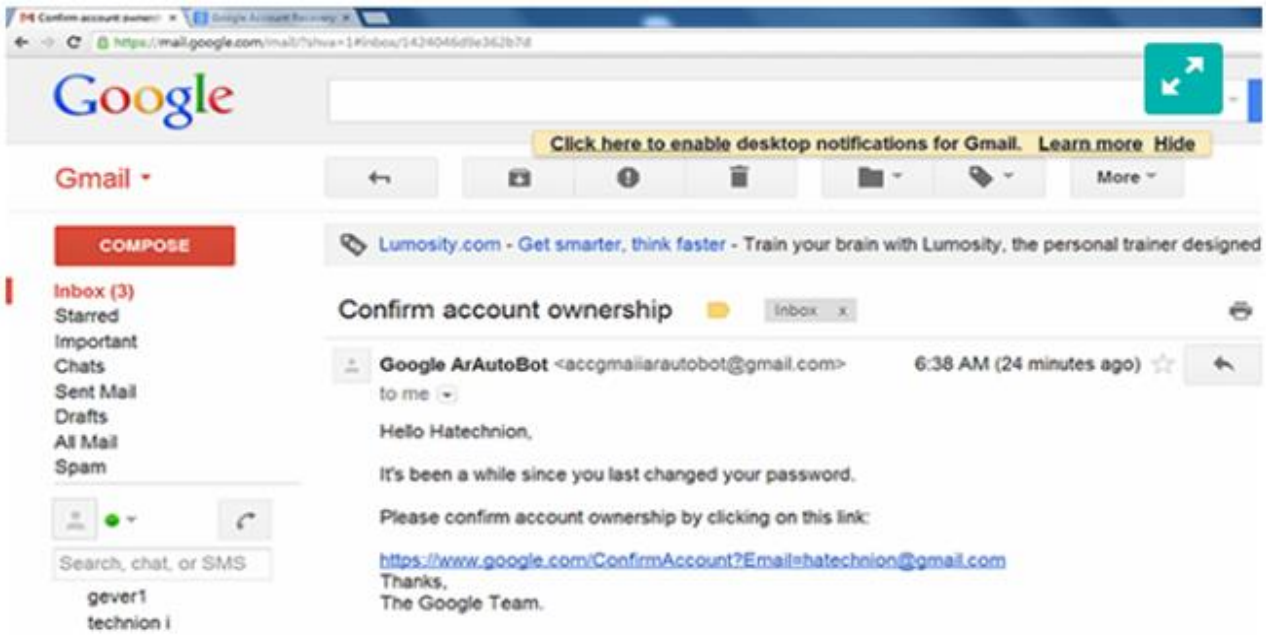

Fig 2:- Example for How Intruder Introduce their Attempt

ISSN No:-2456-2165

The client will get some suspicious emails based on his social media activities, figure depicts that attacker send mail related to the Google account ownership, the mail performed like sent by the original source person but its fake, which is sent by attacker to get clients information By clicking into the link client can share his details to the attacker. So that client has to test the link by adding link to the prediction system.

➢ *Flash Server:*

Flash server collects URL's from the clients and processes those to extract their properties from the internet.

➢ *Internet:*

Internet gives all the properties of legitimate URL's, so by that its helps to classify the URL in to two forms. Properties accessed by the internet are standard and effective once. Those were input to the next step to train our system.

➢ *Module:*

At this stage properties gained by Internet are compared with existing URL and test cases to predict the result about URL, which is legitimate or not. Here in this project we used some algorithms, those are SVM, NeuralNet and decision tree, then compared all those efficiency, based on that we used decision tree algorithm to implement prediction system. Along with these algorithms, Confusion matrix used to classify the test cases. At last predicted results sends to the clients.

C. Algorithm Implemented

Here we implemented decision tree algorithm for efficient and good result

➢ *Decision Tree*

Decision tree algorithm is one of the supervised learning types. Classification and regression problems are solved by decision tree algorithm. In most of data mining problems use decision tree learning. That creates predicted system using the target variables along with several outputs. Problems are solved by tree representation used by decision tree in which class label corresponds to leaf node and internal node corresponds to attributes. Boolean function represents discrete attribute's using this tree.

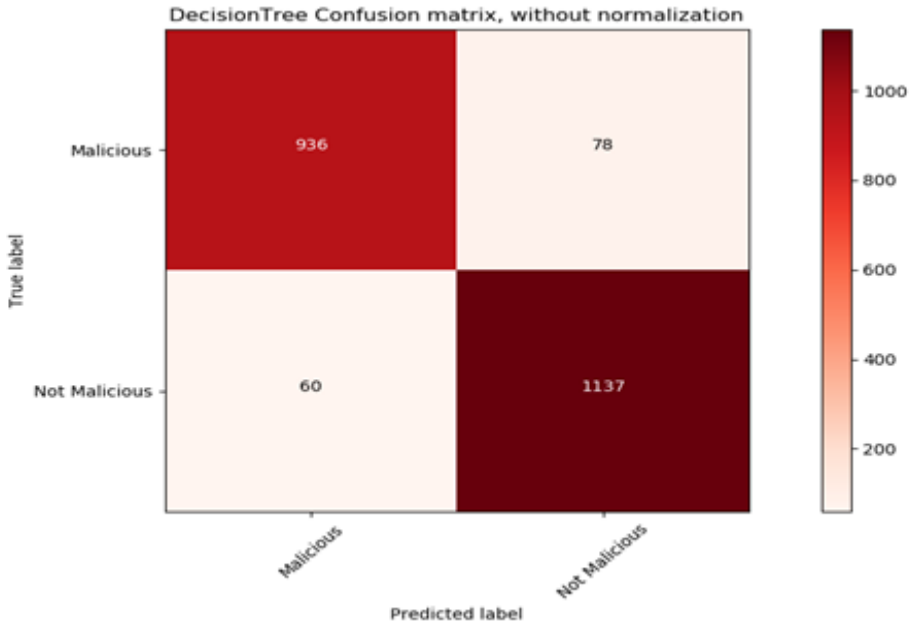

Fig 3:- Confusion Matrix for Decision Tree

The figure 3 depicts the confusion matrix for decision tree algorithm by using some dataset; it classifies the dataset as two attributes such as malicious and not malicious. And its accuracy of the algorithm is 0.937, which has high accuracy compared to other algorithms.

➢ *How Decision Tree Works:*

ID3 is the core algorithm to implement decision tree. ID3 needs to calculate Information gain and Entropy to build decision tree. Naive Bayesian needs entire predictors use Bayes' rule and between predictors there are independence assumptions but decision tree contains entire predictors have dependence assumptions between them.

ISSN No:-2456-2165

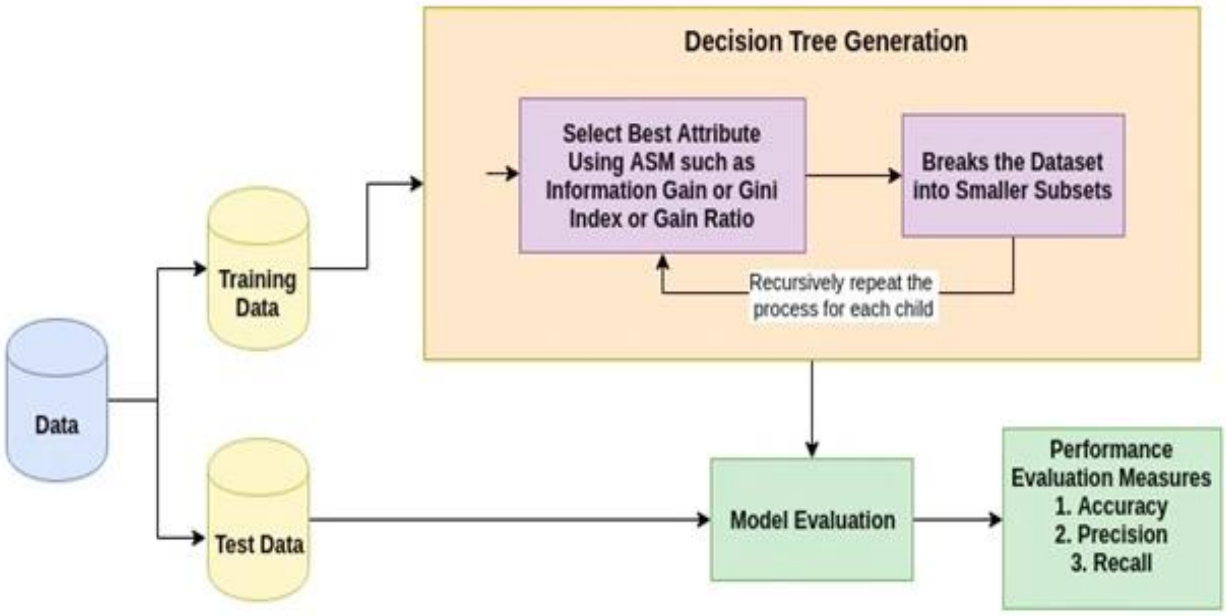

Fig 4:- Decision Tree Workflow

➢ *Entropy*

Top-down approach is built in decision tree from the root node and also includes partitioning the information into subsets that contain instance with homogenous values (similar values). Entropy helps to evaluate the homogeneity of a sample in ID3 algorithm. If entropy is zero means Sample is fully homogeneous and if entropy is one means the sample is equally partitioned it. figure 5 depicts formula for to find entropy.

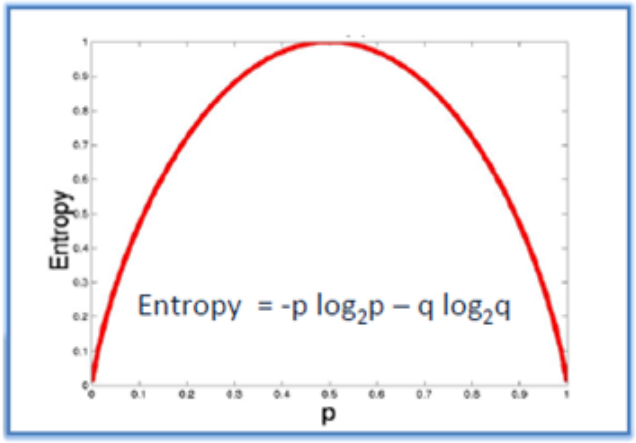

Fig 5:- Calculation of Entropy

There are two types of entropy to build a decision tree, so we need to calculate both these entropy by the help of frequency table. Below two are the formals to calculate the two types of entropy.

$$E(S) = \sum_{i=1}^{c} - p_i \log_2 p_i$$

$$E(T, X) = \sum_{c \in X} P(c) E(c)$$

D. Information Gain

After splitting an attribute there must be some decrease in entropy that is base for information gain. Building decision tree based on finding attribute which has return biggest information gain (i.e. most homogeneous branches).The figure 4 depicts the flow of decision tree algorithm; each steps of algorithm explained next to the figure.

➢ *Steps*:

- Step 1- evaluate entropy of target.
- Step 2- Then evaluate entropy for each branch. Subtract entropy before split from resulting entropy. The result is decrease in entropy (i.e. information gain).

$$Gain(T, X) = Entropy(T) - Entropy(T, X)$$

- Step 3- in this step select attribute with highest information gain as root node, by the branches divide the dataset and then perform same step repeatedly on every branch.
- Step 4-which branch has entropy of zero is a leaf node.
- Step 5-which branch has entropy more than zero sends for further splitting.
- Step 6-repeat algorithm recursively on the branches which is still non leaf, till then all test cases are classified.

IV. EXPERIMENTAL RESULTS

From some recent years, there are so many fake websites those are harmful for the companies and also for individuals. Hence, we have build a model called predict legitimacy of websites, which helps to check whether the links are legitimate or not.

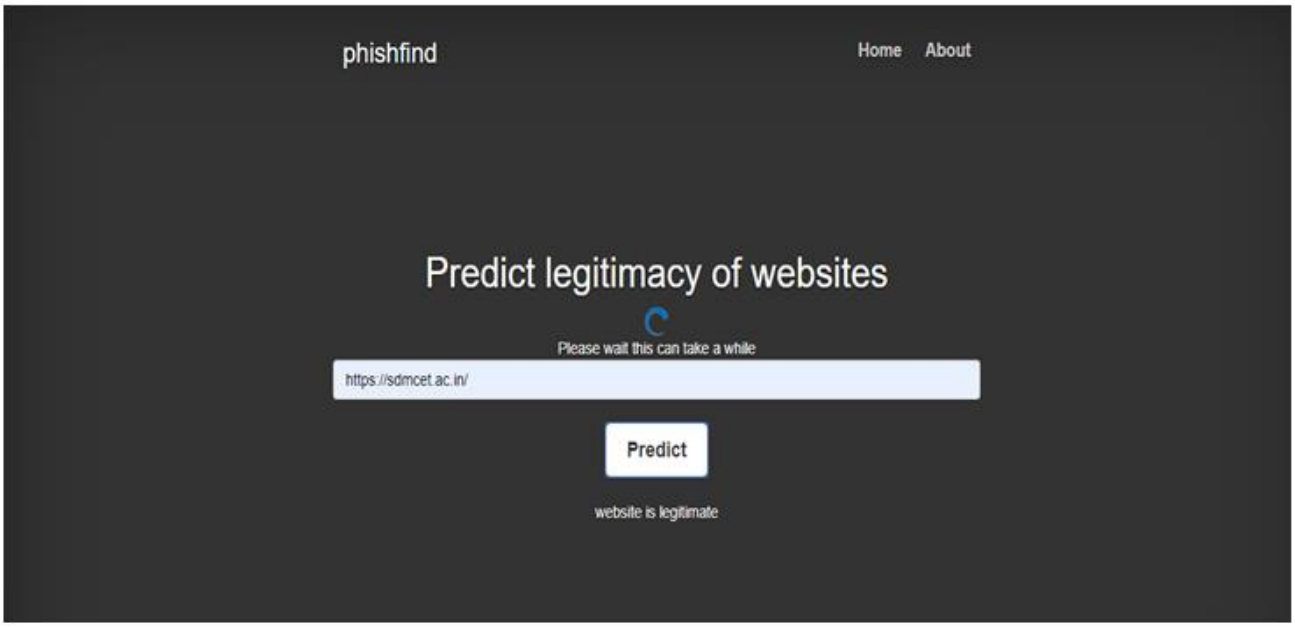

Fig 6:- Predict Legitimate Website

The figure 6 depicts result of legitimate website. Here user copies the link to the application, and then application processes it and provides the appropriate result.

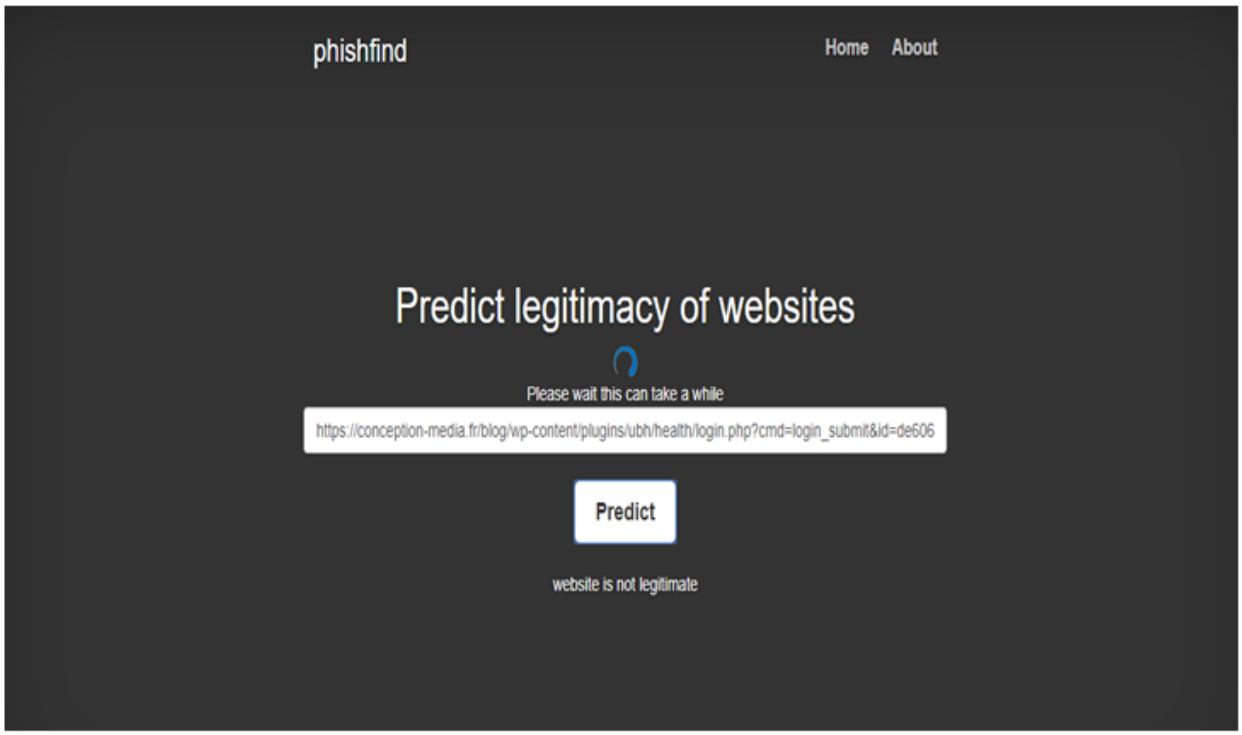

Fig 7:- *Predict Not-Legitimate Website.*

Whenever given links are fake, application processes it and gives result as “not legitimate” which demonstrated in figure 7. Hence, this provides information to the user that the URL may contain any kind of virus.

V. CONCLUSION AND FUTURE ENHANCEMENT

The application developed provides access to all the users to check whether any URL is legitimate or not. Whenever, somebody provides a link through mail or other resources about work, payment or filling of medical or personal details, it becomes important for the user to know whether the link provided is authentic or not. Hence, this application is used to know the status of the URL. If the URL turns out to be false, then any individual can reject the link by not downloading it or filling the details in it. This saves the sensitive data of any individual from leaking out from the system. Hence, the system data is preserved. The manual prevention of APT's is proposed here but automation preventions will see in future enhancement.

The application checks whether the URL given is legitimate or not. But in future, the application can be expanded with checking of contents from all the sites for legitimacy. For example, some of the contents given in the sites may be false. Hence, this application added with future enhancement will let user to distinguish about the fake and real information. And the manual prevention of APT's is proposed in our project but automation preventions will see in future implementation.

REFERENCES

[1]. Bencsáth, B., Pék, G., Buttyán, L., & Felegyhazi, M. (2012). The cousins of stuxnet:Duqu,flame, and gauss.Future Internet ,4(4),971-1003. doi:10.3390/fi4040971
[2]. FireEye, Inc. (2016).FireEye threat intelligence follow the money: Dissecting the operations of FIN6.Retrieved May 6, 2016, from
[3]. Langner, R. (2011). Stuxnet: Dissectinga Cyberwarfare weapon. IEEE Security & Privacy Magazine, 9(3), 49 – 51. doi:10.1109/msp.2011.67
[4]. Marschalek, M., Kimayong, P., & Gong, F. (2014, November).Point of sale (POS) Malware revisited. Retrieved May 6, 2016, from Cyphort.
[5]. McAfee, Inc. (2011).Combating Advanced Persistent Threats. Retrieved May6,2016,From [6] Moon, D., Im, H., Lee, J. D., & Park, J. H. (2014). MLDS: multi-layer defense system for preventing advanced persistent threats.
[6]. Pingree, L., & MacDonald, N. (2012, January 18).Best Practices for Mitigating Advanced Persistent Threats (Rep.). Retrieved April 21, 2016, from Trend Micro, Inc.
[7]. Symantec, Inc. (2010).What is a Zero-Day vulnerability? Retrieved May 10, 2016, From PCTools.
[8]. Wangen, G. (2015). The Role of Malware in Reported Cyber Espionage: A Review of the Impact and Mechanism. Information, 6(2), 183-211. doi:10.3390/info6020183