International Journal of Modern Engineering Research (IJMER)
www.ijmer.com Vol.3, Issue.2, March-April. 2013 pp-1245-1252 ISSN: 2249-6645

Anomaly Detection Using Generic Machine Learning Approach With a Case Study of Awareness

Goverdhan Reddy Jidiga, [1] Dr. P Sammulal[2]
[1]*Research Scholar, JNTU & Lecturer in CSE, Department of Technical Education, Govt. of Andhra Pradesh, India*
[2]*Senior Assistant Professor, JNTUCEJ, JNT University Hyderabad, Andhra Pradesh, India*

Abstract: *Security of computer systems and information in flow is essential to acceptance for every network user utilities Now the standalone computer and internets are exposed to an increasing number of security threats with new types of attacks continuously appearing. For this to develop a robust, flexible and adaptive security oriented approaches is a severe challenge. In this context, anomaly based intrusion detection technique is an advanced accurate technique to protect data stored at target systems and while flow in the networks against malicious activities. Anomaly detection is an area of information security that has received much attention in recent years. So in this paper we are going to elaborate a latest techniques available in machine learning approach applied to anomaly detection which are used to thwarts the latest attacks like cyber based attacks and malware infections. Finally a case study is discussed on latest cyber attacks phased by top web domains and countries in the world motivated by a traditional security ethic are called E-Awareness.*

Keywords: *Anomaly, Cyber attacks, Machine learning, Malware, Phishing,*

I. INTRODUCTION

Every day the cyber criminals are invading countless homes and offices across the nation not by breaking down windows and doors, but by breaking into laptops, personal computers, and wireless devices via hacks and bits of malicious code. For this billions of dollars are lost every year repairing systems hit by such attacks. Some take down vital systems, disrupting and sometimes disabling the work of hospitals, banks, other services around the world. Today, these computer intrusion cases like counterterrorism, counterintelligence, and criminals are the paramount priorities of our cyber program because of their potential relationship to national security. The solutions are there, but not terminate permanently so only case that gather vital intelligence that helps us identify the cyber crimes that are most dangerous to our national security and to our economy.

Anomaly detection is type of Intrusion detection and the Intrusion detection [29] defined as the process of monitoring the anomaly based events occurring in a computer system or network and analyzing them for signs of intrusions, defined as attempts to bypass the standard security mechanisms of a clean computer or network that are compromise the confidentiality of the valuable data, data integrity, availability and access control of information sources as well as resources. Intrusion detection system (IDS) [15] is a combination of software and hardware that attempts to perform intrusion detection protection to normal users and system resources from information security threats. Computer security analysts use intrusion detection systems to assist them in maintaining computer system security. There were numerous attacks on software systems result in a process execution or human coding mistakes deviating from its normal behavior[2] , all these prominent examples include a malware related code injection attacks on internet server's processes and with resulting from buffer overflow and format string vulnerabilities. Up to now we have seen significant amount of research to detect such attacks through monitoring the behavior of the suspected process and comparing that behavior to a model [3] consist of normal behavior collected from past experiences. These are also called anomaly detection techniques because in compare to signature based detection which deviates from the normal behavior are taken as indications of anomalies.

II. ANOMALY DETECTION

2.1 Introduction

In this the basic unit of finding abnormal behavior is identified as an anomaly. Anomaly [8, 29] is a pattern in the data that does not conform to the expected behavior, depending on the nature of input data collected from profiles that represent normal behavior of users, applications, hosts, networks, and detecting attacks as significant deviations from this profile. Anomaly detection is monitors program executions and detects anomalous program behaviors through reverse analysis of executable program including a critical behavior monitoring points can be extracted from binary code sequences [2] and memory space. Most of the available intrusion detection systems are predominantly signature based, so such systems are not used to find the frequent rule based attacks, unknown attacks and updates. The existing systems even if it is designed by traditional and advanced anomaly detection techniques are not observing real world anomalies like emerging cyber threats, cyber intrusions, credit card frauds...etc. The anomalies occur relatively infrequently and their consequences can be quite dramatic, negative sense in the running of applications

2.2 Why Anomaly Detection

The security breaches are very common now in the society and organizations fail to take effective measures, with the legal action by authorities over breaches of and devastating damage to brands, reputation and customer confidence. When it comes to new technologies organizations have needed to move quickly, but they are not responding fast. Now smart

phones and latest gadgets were primarily used by everyone, so in the same way risks related to mobile devices, social media and in critical infrastructure are worst[21]. The business and personal use they have forced organizations to urgently implement policies that address the risks associated with an evolving array of emerging technologies.

Also the organizations are protecting how they guard to their data of the employees and their customers like service based organizations in the world phasing number of problems from attacks and threats with latest criminal mind showing their illegal operations. Today cyber attacks are common in the public banking sector, health organizations, defense, and service sector , so organizations are need to give training and guidelines, policy adjustments, stepping up awareness programs. So our aim is to prepare for an effective solutions working online and on the fly counter action is required to avoid the cyber criminals, viruses, malware and botnets shown. The security teams and experts must work round the clock to actively manage the risks created by threats in different top rated domains like business, online-shopping, social media including comprehensive policies and effective security controls. Now experts need to not only consider how they can occur and use powerful analytics to detect security events but also realize to aware of dynamic threats caused by malicious events.

Area/ Domain / Threat (Anomaly)	Stable Computers	Networks &Comm'n	Banking Sector	Health and Insurance	Social Network'g	Service sector orgs	Defense	World Economic	Supply-Chain Marketing	Cloud Computing	ICT/IT Sector
Cyber fraud	•	✓	✓	✓	✓	✓	✓	✓	✓	•	o
Malware	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Insecure API	•	✓	✓	✓	•	✓	✓	•	•	✓	✓
Wanted assaults	✓	✓	✓	✓	✓	✓	✓	o	•	o	✓
Organized crime	•	✓	✓	✓	✓	✓	✓	•	✓	✓	✓
Cyber war	•	✓	•	•	✓	✓	✓	✓	o	o	•
Cyber spying	✓	✓	✓	✓	✓	✓	✓	•	o	✓	✓
Nation states	•	✓	✓	•	✓	•	✓	✓	•	•	o
Big Data	•	•	o	✓	o	•	•	•	•	✓	o
Program source threats	✓	✓	•	•	✓	✓	•	•	•	✓	✓
Hacking as a service	•	✓	✓	✓	✓	✓	✓	o	✓	•	✓
Device Target (BYOD)	✓	✓	•	✓	✓	o	•	•	•	•	✓
Social Crime	•	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hactivism	✓	✓	✓	✓	✓	✓	✓	✓	✓	•	•
Fake certificates	•	✓	✓	✓	✓	✓	✓	•	✓	✓	•
Mobile Malware	✓	✓	•	•	•	✓	•	•	•	•	✓
Ransomware	✓	✓	✓	✓	✓	✓	o	✓	✓	o	✓
Cyber Miasma	•	•	•	•	•	•	•	✓	•	•	✓

• Low o Medium ✓ High

Fig.1. Threats Vs Area/Domain in 2012

According to information security survey [21, 22, 23] up to 2012 the threats and malware infections on the top list found in the world is classified according to area or domain as following Figure-1. In this there are many critical infrastructure top web domains are affected by different threats and their impact low, medium, and high. To overcome these threats we use the anomaly detection with usage of machine learning approaches [29]. The main uses for Anomaly Detection are detect precedent attack behavior, zero day attack detection, intrusion model, insider threat detection, situational awareness and validate or assist with signature data. The major advantages of anomaly detection are network intrusion detection, insurance or credit card fraud detection, healthcare informatics or medical diagnostics, industrial spoil detection, video surveillance or image processing, novel topic detection in text mining and so on. Anomaly detection (AD) systems have some advantages [27, 29]. First AD have the capability to detect insider attacks like someone using a stolen account starts performing actions that are outside the normal user profile, an anomaly detection system generates an alarm. Second, AD is based on customized profiles, it is very difficult for an attacker to know with certainty what activity it can carry out without setting off an alarm. Third, an anomaly detection system has the ability to detect previously unknown attacks.

III. RELATED WORK

Most of the anomaly intrusion detection systems are signature based and fundamental statistics or knowledge based, but these are all suitable in some applications and not suitable today in advanced technical concepts. Now we discussed some related work on old and new one is based on machine learning discussed in next paragraph. Anderson is the first person elaborated the intrusion concept in security and he developed model [1] threats are classified as external penetrations, internal penetrations, and misfeasance. The Anderson model [5] is good initially, but now it is not suitable. Denning proposed several models for Intrusion Detection System (IDS) development based on statistics, Markov chains, time-series, etc [4]. In Denning model, user's behavior that deviates sufficiently from the normal behavior is considered anomalous. Stanford Research Institute developed an Intrusion Detection Expert System (IDES) that continuously monitored user behavior and detected suspicious events. Later SRI developed an improved version of IDES called the Next-Generation Intrusion Detection Expert System (NIDES) [05] that could operate in real time for continuous monitoring of user activity. NIDES enable the system to compare the current activities of the user/system/network with the audited intrusion detection variables stored in the profile and then raise an alarm if the current activity is sufficiently far from the stored audited activity. A statistical anomaly based IDS were proposed by Haystack [29], which used both user and group based anomaly detection

strategies. In this system, a range of values were considered normal for each attribute and during a session if an attribute fell outside the normal range then an alarm raised. It was designed to detect six types of intrusions: attempted break-ins by unauthorized users, masquerade attacks, penetration of the security control system, leakage, denial of service, and malicious use. SNORT is an open source network intrusion detection and prevention system (NIDPS) developed by Sourcefire. In 1996, Forrest proposed an analogy between the human immune system and intrusion detection that involved analyzing a program's system call sequences to build a normal profile [10], if the sequences deviated from the normal sequence profile then it considered as an attack. The system they developed was only used offline using previously collected data and used a quite simple table-lookup algorithm to learn the profiles of programs.

Machine learning based work: In 2000, Valdes [11] developed an anomaly based intrusion detection system that employed naïve Bayesian network to perform intrusion detecting on traffic bursts. In 2003, Kruegel [26] proposed a multisensory fusion approach using Bayesian classifier for classification and suppression of false alarms that the outputs of different IDS sensors were aggregated to produce single alarm. In the same year, Shyu [12] proposed an anomaly based intrusion detection scheme using principal components analysis (PCA), where PCA was applied to reduce the dimensionality of the audit data and arrive at a classifier that is a function of the principal components. In [19,20] proposed an anomaly based intrusion detection using hidden Markov models that computes the sample likelihood of an observed sequence using the forward or backward algorithm for identifying anomalous behavior from normal behaviors. Lee [7, 13] proposed classification based anomaly detection using inductive rules to characterize sequences occurring in normal data. In 2000, Dickerson [14] developed the Fuzzy Intrusion Recognition Engine using fuzzy logic that process the network input data and generate fuzzy sets. So therefore, the primary and most important challenge is we needs to be develop the on the fly countermeasures and effective strategies to reduce the high rate of false alarms.

IV. MACHINE LEARNING

In this paper we concentrated on machine learning techniques , because machine learning approaches use strong statistical foundations to enhancing the dynamic and accurate learning that gives better accuracy, small false alarm rates, learned detectors use a more compact representation, possible performance improvements, ability to detect novelty, protection against zero-day exploits, faster incident response times.etc. Machine Learning systems [17, 27] offer unparallel flexibility in dealing with evolving input in a variety of applications, such as intrusion detection systems and spam e-mail filtering. However, machine learning algorithms themselves can be a target of attack by a malicious adversary. In the machine learning different novel contributions of techniques include taxonomy of different types of attacks on systems, a variety of defenses against those attacks.

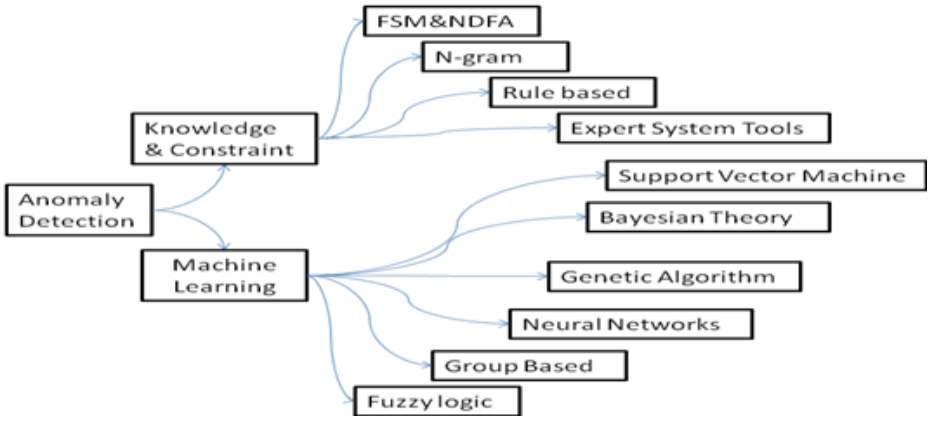

Fig.2. Taxonomy of Anomaly Detection based on Machine learning.

In the fig.2 the anomaly detection taxonomy[8,15] is given, it is based on classification of anomaly detection which is purely based foundational work done past authors of intrusion detection systems models and today performance based machine learning approaches. In this taxonomy the anomaly detection is based on machine learning and data mining approaches. The machine learning approaches use strong statistical foundations to enhancing the dynamic and accurate learning that gives better accuracy, small false alarm rates.

4.1 Proposed Model

In fig.3 the machine learning based AD (Anomaly Detection) is used and prototype is given with preprocessing data. In AD (Anomaly Detection) prototype model the audit data collection module is used in the data collection phase. The data collected in this phase is analyzed by the anomaly detection algorithm to find traces of suspicious activity. The source of the data can be host/network activity logs, command-based logs, application-based logs, etc. audit data in intrusion detection systems store the audit data either indefinitely or for a sufficiently long time for later reference. The volume of data is often exceedingly large, so persistent database is maintained.

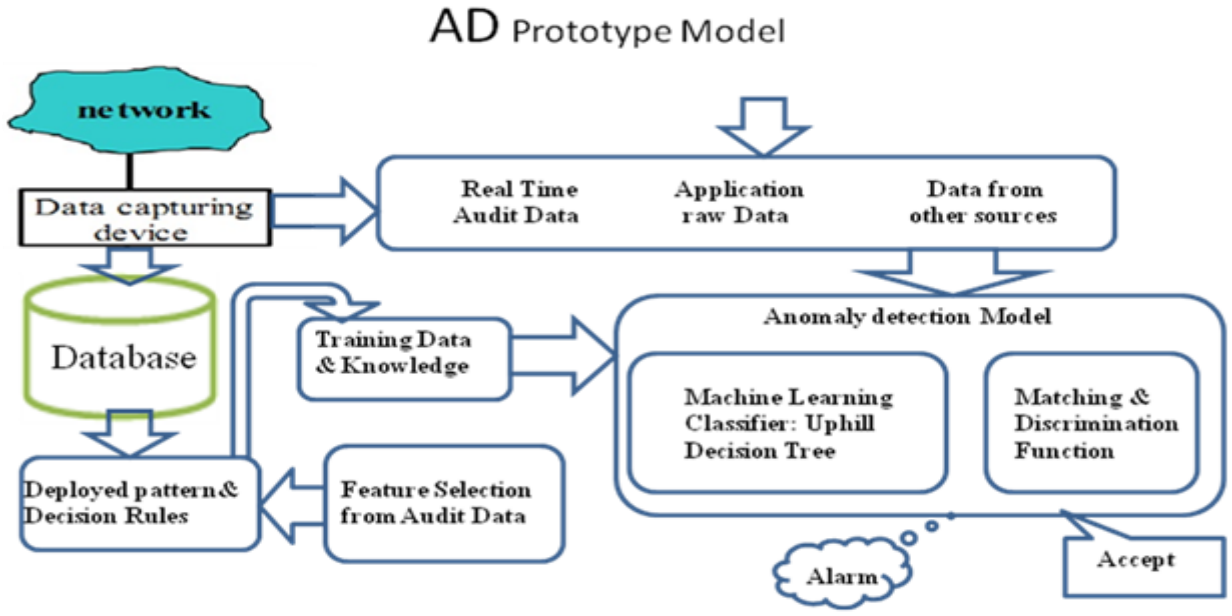

Fig.3. Anomaly Detection (AD) Prototype Model with Machine learning.

In Anomaly detection model the processing is here the algorithms to detect suspicious activities are implemented. Algorithms for the analysis and detection of intrusions have been traditionally classified into three broad categories: signature (or misuse) detection, anomaly detection and hybrid (or compound) detection. The configuration data is extracted from network database and user deployed rules framed from information that is pertinent to the operation of the intrusion detection system itself such as information on how and when to collect audit data, how to respond to intrusions, etc. Due to the complex and dynamic properties of intrusion behaviors, machine learning and data mining methods have been generally employed to optimize the performance of intrusion detection systems to finding specific point anomalies or range anomalies at moment of time. We give an efficient algorithm for provably learning uphill decision tree with extension adornments of existing multi-way decision tree algorithm. In fig.3 the machine learning is decision tree algorithm is considered first and later it is extended to uphill decision tree also called regression tree. The processing element must frequently store intermediate results such as information about partially fulfilled anomalies. The model contains a logic taken from uphill decision tree detect the anomalies by raising notified alarms.

4.2 Proposed Machine Learning Algorithm : A Uphill Decision Tree (UDT)

The Decision tree (DT) learning [18, 29] is a type of machine learning algorithm used in many application of information security in previous research. The decision tree (DT) is very powerful and popular data mining algorithm for decision-making and classification problems. It has been using in many real life applications like medical diagnosis, radar signal classification, weather prediction, credit approval, and fraud detection etc. DT can be constructed from large volume of dataset with many attributes, because the tree size is independent of the dataset size. A decision tree has three main components: nodes, leaves, and edges. Each node is labeled with an attribute by which the data is to be partitioned. Each node has a number of edges, which are labeled according to possible values of the attribute. An edge connects either two nodes or a node and a leaf. Today it is olden and not effecting in current cyber attacks. So the extension of this is an uphill decision tree.

A decision tree with real values at the leaves is called an uphill decision tree if the values on the leaves are non-decreasing in order from left to right. An Uphill decision tree is similarly a tree structured solution in which a constant or a relatively simple regression model is fitted to the data in each partition. In this algorithm we considered the data collected for e-mail to filtering whether it is phishing or spam.

Phishing [21] is a type of Internet fraud deployed to steal confidential financial information that includes theft of net banking passwords, corporate secrets, credit card numbers, financial status, bank account details and other valuable information and spam is anonymous, unsolicited bulk email diverting the cybercitizen's minds to use their services and products etc. phishing is also type of spam.

The total number of phishing emails and spam are increasing day to day, up to 2012 the top domains and areas in the world suffering with cyber threats like DOS, DDOS,SQL Injection, spamming attacks , phishing attacks and others. In that most of them are spam , phishing attacks affected on online banking, Online purchasing (PayPal, Amazon, eBay, etc.), Social media (Face book, Twitter, blogs, etc.) in all corners.

The semantics of e-mail are like domain, class, frequency, link, URL, IP address, script, validation, port address, dot, images, no. of ports valid or not, link valid or not, mismatching ..Etc available. Based on that we can estimate that the mail or websites are legitimate or phishing by using a pre determined set of rules designed in the construction of uphill decision tree.

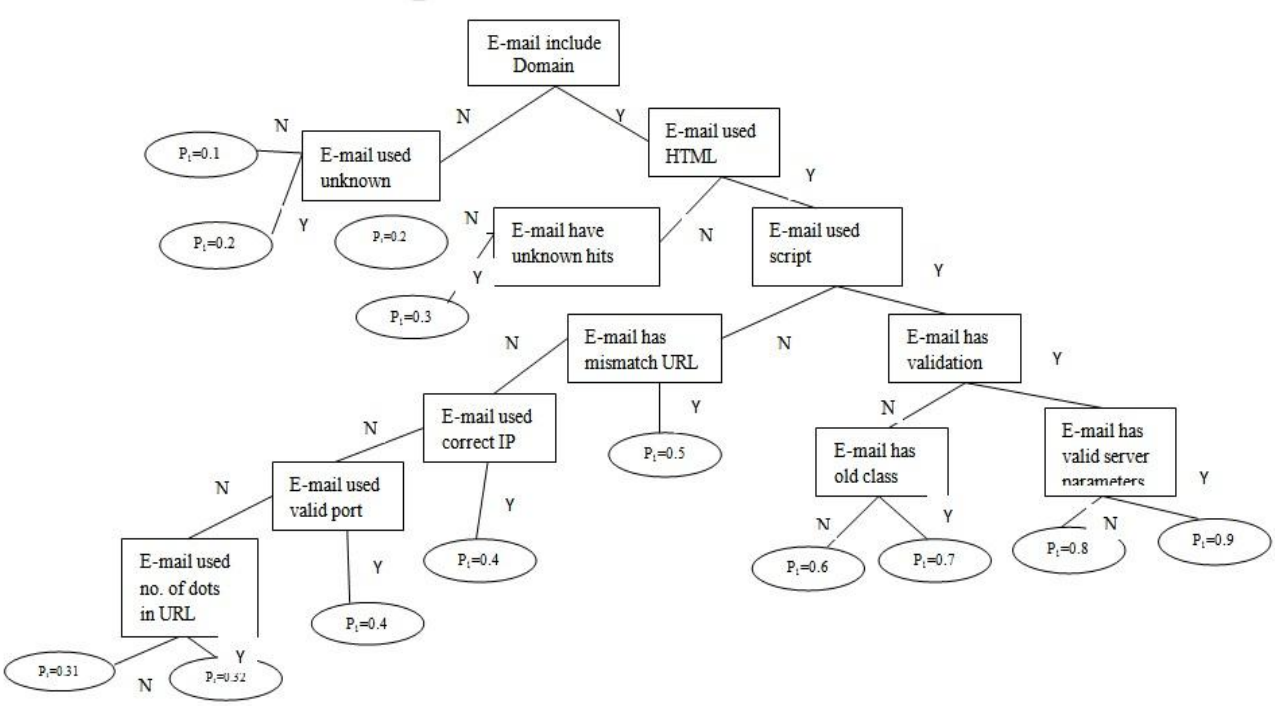

Fig.4. Uphill Decision Tree for e-mail verification to estimate the Phishing attack.

In fig.4 we have applied a concept of uphill decision tree (statistics are assumptions only) for e-mail verification whether the e-mail is come from legitimate organizations or not. The mail semantics are compiled by tree one by one and find the some unknown semantics are encountered that we compare with original semantics of e-mail.
In this we take an example-1 of e-mails is registrar@jntuh.ac.in, from jntuh.ac.in/new/academic/contacts_us.html, in JNTU.

Here if (domain=TRUE&e-mail-has=HTML&script=Java_script&Validation=TRUE&server=authenticate)
Then P (e-mail Semantics) = "90%". (This value depending on other semantics also)

In this e-mail, if all semantics are correct and verified by decision tree including URL, no. of dots in URL, IP address and port of application then we can probably identify that mail come from authorized party.

Example-2: VISA card related mail from VISA organization contains unknown hits like "dear valued customer" shown in fig.5, But in original mail from bank is not contains semantics like "dear valued customer".

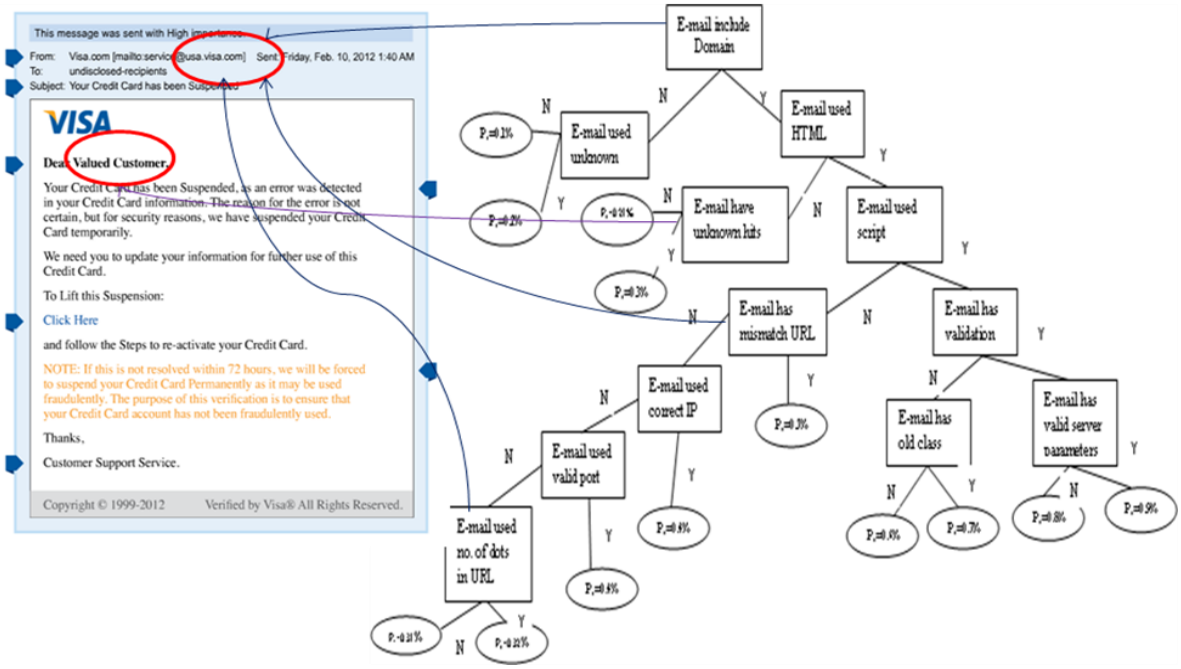

Fig.5. Phishing attack based e-mail verification came from VISA

Here if (domain=TRUE &e-mail-has="unknown hits") then P (e-mail Semantics) = "20%". (This value depending on other semantics also)

In this e-mail, first semantic is correct and verified by decision tree and semantic at node-3 has unknown hits may not included in semantics of VISA mail including URL, no. of dots in URL, IP address and port of application then we can probably identify that mail come from unauthorized party.

4.3 Ethical Solution to Phishing : E-Awareness and Case Study

The basic E-security awareness [23] is a process of keeping people in continuous attention of security to save information. In the abstract we clearly mentioned e-awareness also a part information security ethics to thwart a most vulnerabilities as "security awareness is better than prevention and prevention is better than detection". This was an ethic concept applied in all kinds of human life applications to survive in the nature. The people who are expert in security aspects to thwart the security deficiencies, eligible to train all users of information technology to identify and report the all kind of suspicious activities in their electronic environment. Now it is essential that each of us take responsibility and understand our role in securing cyberspace.

ACM Report[28] given the countermeasures on phishing is creating awareness and train end users to proactively recognize and avoid phishing attacks (ethical and very popular approach).The solutions are motivating people to be secure, micro games designed to teach people about phishing and embedded training.

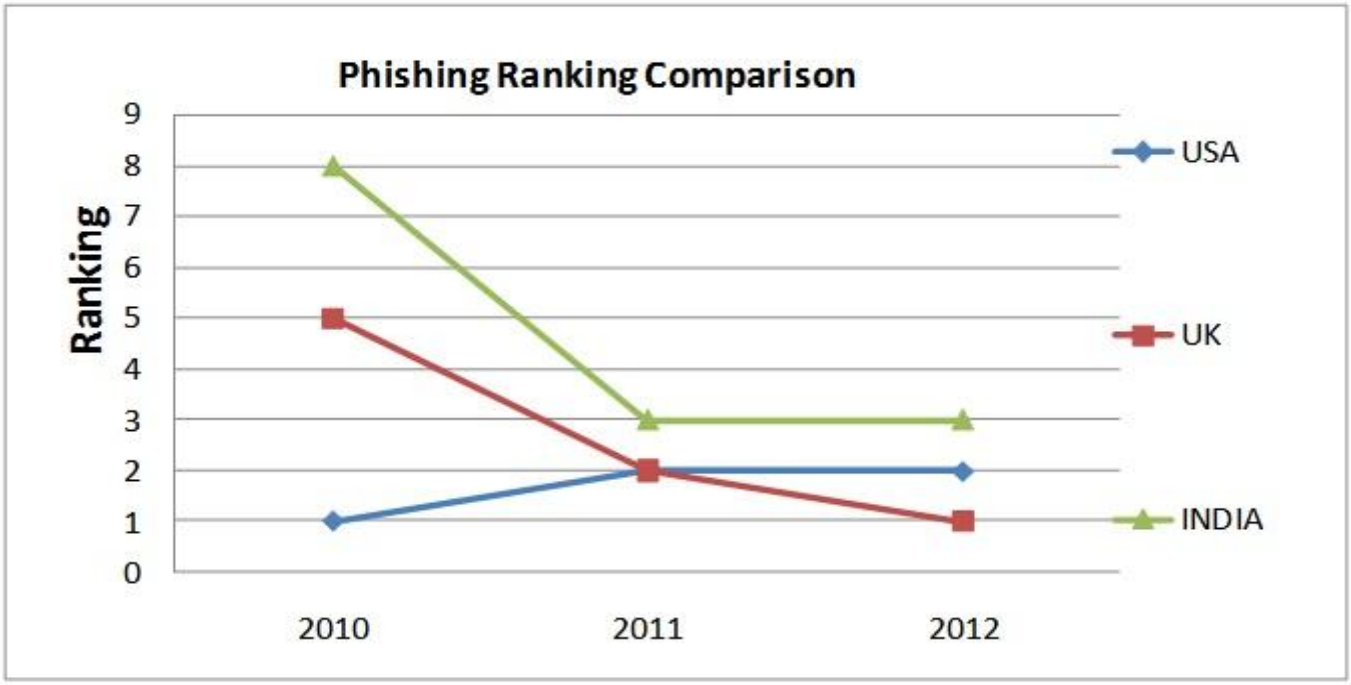

Fig.6. Phishing attack Ranking

From 2010 to 2013 USA ranking is going to down shown in fig.6, because the public and private partnership security awareness months [21, 23, 24] programs conducted by hundreds of organizations like

1) National Cyber-Security Alliance, FBI
2) U.S. Department of Homeland Security
3) DHS with RSA-EMC, Google, McAfee, Microsoft, Symantec, ADP, AT&T, Cisco, IBM
4) CERT, FTC, NCL.

The DHS of USA taking an active involvement in the control of all types of cyber threats by creating awareness, responsibility, co-operation, policy enforcement, and also motivating people to learn ethics.

The reasons to reduce the phishing in USA:

1) They are conducting fast track E-security awareness programs by monthly in different states.
2) Most of government and private organizations believing that "Awareness is only an impressive and vital element of the phishing combat".
3) USA promoting awareness, education, research on addressing the latest challenges well in advance.
4) Regular recommendations to update policies, guidelines, and risk assessment.
5) USA anti phishing group suggesting to use look ahead technology to avoid loss.
6) Respond rate is high due to education and motivation.
7) In 2012 USA conducted 120 awareness programs conducted approximately.
8) Regular alerts and reports.

The reasons to increase the phishing attacks in UK:

1) UK banking and financial institution customers use heavy online money transactions.
2) 84% people of UK using internet daily at least once and create breaches.
3) In UK most of the people use the mobile internet transaction leads to increasing phishing and also malware infection.
4) Insider impersonation of the organization.
5) Lack of coordination among the consumers and financial institutions to report latest trends in phishing.
6) Latest brands in the market

What are the reasons for increasing phishing and spam [22] in India are :

1) Lack of awareness, education and responsibility.
2) Lack of use of new technology and using of poor technology.
3) High unemployment, illiterate, ignorance, population and competition in the market.
4) Lack of government support, policy constraints, coordination, law.
5) Greediness in earning of easy money.
6) May fast economy development and technology using.
7) Use of pirated software.
8) Rate of using internet.
9) Huge growth in the usage of Mobiles.
10) System sharing.

V. CONCLUSION

In this paper we accepted only a machine learning approach to anomaly detection and applying to phishing attacks. Now phishing will persist in any electronic medium pursue a problem that can never truly be solved. In this nest better we can work on always to preventing, detecting, and responding to this e-awareness. Finally in this paper we present a case study on phishing attack based on the awareness model and today the machine learning is only approach encouraged by well known scientists in the field of security. So this will give concepts and motivates to you a do further research and also hope that this work to be true at our knowledge.

REFERENCES

[1] J.P.Anderson,"Computer security threat monitoring and surveillance," James P Anderson Co.,Fort Washington,Pennsylvania, USA, Technical Report 98–17, April 1980.

[2] H. H. Feng, Oleg M. Kolesnikov, P. Fogla, Wenke Lee, and Weibo Gong, "Anomaly Detection Using Call Stack Information"IEEE Symposium on Security and Privacy'2003, CA, Issue Date: 11-14 May2003 pp: 62-75 ISSN: 1081-6011 Print ISBN: 0-7695-1940-7.

[3] D. Wagner and D. Dean, "Intrusion Detection via Static Analysis", IEEE Symposium on Security and Privacy, Oakland, CA, 2001.

[4] D. E. Denning "An intrusion detection model" In IEEE Transactions on Software Engineering, CA,1987. IEEE Computer Society Press.

[5] D. Anderson, T . Frivold, and A.V aldes."Next-generation intrusion detection expert system (NIDES): A summary" Technical Report SRI–CSL–95–07,Computer Science Laboratory,SRI International, May 1995.

[6] Mukkamala,J.Gagnon,andS. Jajodia."Integrating data mining techniques with intrusion detection methods" Research Advances in Information Systems Security, Kluwer Publishers, Boston, MA. 33-46,2000.

[7] W.Lee, ChanP.K, Eskin, E WeiFan, Miller M. S.Zhang "Realtime datamining based intrusion detection" IEEE DARPA information Conference 2001,DISCEX'01,Proceedings IssueDate: 2001 page(s):89-100vol.1 12 Jun2001-14 Jun 2001 Print ISBN:0-7695-1212-7.

[8] S.Axelsson,"Intrusion Detection Systems: A Survey and Taxonomy," Chalmers University, Technical Report 99-15,March 2000.

[9] S.E.Smaha,"Haystack:An Intrusion Detection System," in IEEE Fourth Aerospace Computer Security Applications Conference, Orlando, FL, 1988, pp. 37 – 44.

[10] S. Forrest, S. A. Hofmeyr, A. Somayaji, and T. A. Longstaff, "A Sense of Self for Unix Processes," in IEEE Symposium on Research in Security and Privacy, Oakland, CA, USA, 1996, pp. 120--128.

[11] A. Valdes and K. Skinner, "Adaptive Model-based Monitoring for Cyber Attack Detection," in Recent Advances in Intrusion Detection Toulouse, France, 2000, pp. 80-92.

[12] M.L.Shyu,S.C.Chen,K.Sarinnapakorn, and L.Chang,"A Novel Anomaly Detection Scheme Based on Principal Component Classifier," in IEEE Foundations and New Directions of DataMining Workshop, Florida, USA, 2003, pp. 172-179.

[13] W. Lee and S. J. Stolfo, "Data mining approaches for intrusion detection," in 7th USENIX Security Symposium (SECURITY-98), Berkeley, CA, USA, 1998, pp. 79--94.

[14] J.E.Dickerson and J.A.Dickerson,"Fuzzy network profiling for intrusion detection",in 19th Intern'l Conference of the North American Fuzzy Information Processing Society(NAFIPS),Atlanta, 2000, pp. 301 – 306.

[15] L.Ertoz, E.Eilertson, A.Lazarevic, P.N. Tan, V. Kumar, J. Srivastava, and P. Dokas, "The MINDS - Minnesota Intrusion Detection System", in Next Generation Data Mining Boston: MIT Press, 2004.

[16] S.Mukkamala, G.I.Janoski, and A.H.Sung."Intrusion Detection Using Support Vector Machines",Proceedings of the High Performance Computing Symposium- HPC 2002, pp 178-183, San Diego, April 2002.

[17] T. Lane and C. E. Brodley. "An Application of Machine Learning to Anomaly Detection", Proceedings of the 20th National Information Systems Security Conference, pp 366-377, Baltimore, MD. Oct. 1997.

[18] Quinlan, J. Ross, "Induction of Decision Trees," Machine Learning, 1:81{106, 1986. Reprinted in Shavlik, J. and Dietterich, T.,Readings in Machine Learning, San Francisco: Morgan Kaufmann, 1990, pp. 57-69.

[19] Ghahramani Z,"An introduction to hidden markov models and bayesian networks". HMM: applications in computer vision, pages 9–42, 2002.

[20] HaiTao H,XiaoNan L,"A novel HMM-based approach to anomaly detection", Journal of Information and Computational Science 1 (3) (2004) 91–94.

[21] http://www.antiphishing.org/

[22] http://india.emc.com

[23] http://www.rsa.com

[24] http://www.cisco.com

[25] A. Ghosh and A. Schwartzbard, " A study in using neural networks for anomaly and misuse detection", 8th USENIX Security Symposium, pp. 141-151, 1999.

[26] C. Kruegel, D. Mutz, W. Robertson, and F. Valeur, "Bayesian Event Classification for Intrusion Detection," in 19th Annual Computer Security Applications Conference, Las Vegas, NV, 2003.

[27] L.Breiman,"Random Forests,"Machine Learning,vol. 45, pp. 5-32, 2001.

[28] By Jason Hong, "article on The State of Phishing Attacks" Communications of the ACM, Vol. 55 No. 1, Pages 74-81.

[29] A. Patcha, J-M. Park, "An Overview of Anomaly Detection Techniques: Existing Solutions and Latest Technological Trends", Computer Networks(2007)

[30] M. Roesch, "Snort - Lightweight Intrusion Detection for Networks " in Proceedings of the 13th USENIX conference on System administration Seattle, Washington 1999 pp. 229-238.