Data Shield Algorithm (DSA) for Security against Phishing Attacks

Ram Avtar[1], Bhumica Verma[2] and Ajay Jangra[3]

[1]Electrical Engg. Department, [2,3]CSE Department,

UIET, Kurukshatra University, Kurukshetra, INDIA

[1] ramavtar.jaswal@gmail.com, [2] talk2bhumica@gmail.com, [3] er_jangra@yahoo.co.in

Abstract: *The World Wide Web provides every internet citizen with voluminous and heterogeneous data. Therefore, it becomes an essential to mine this available data to make it presentable, useful, and pertinent to a particular problem. Web mining deals with the extraction of these interesting patterns and developing useful abstracts from diversified sources. To improve the security of Web services one would increase its Quality, hence pre-fetching scheme are opted which increases network traffic and Web server's load that is used by DDoS Attack. The DDoS attack mostly targets the computer network's bandwidth or connectivity. Here the paper focuses on network based DDoS attacks especially the phishing Attacks. This paper concentrates on security against such attack and an novel "Anti-phishing System" is designed with a view to provide an instant, automatic, comprehensive system level solution to perform webpage authentication and webpage detection against phishing and named it as Data Shield algorithm. Data Shield algorithm detects both the known and the unknown phishing attacks. This developed algorithm is 'Real-time Light weighted Anti-Phishing Algorithm' that can detect and prevent the users from phishing attacks in the real time.*

Keywords: *Network Security, DDoS Attacks, Phishing Attacks, Anti-Phishing system, DataShield Algorithm.*

1. Introduction

The term Denial of service (DoS) attack is used to get any specific information where a malicious user create obstacle and limits the authorized users to access network services by using the resources of the victim system [1]. These attacks don't necessarily damage data directly or permanently, but they intentionally compromise the availability of the resources. The DoS attacker [2] creates network congestion by generating a large volume of traffic in the area of the targeting system. The size of the caused overload is enough to prevent any packet from reaching its destination. In a DDoS attack [3], the aggregation of the attacking traffic can be tremendous compared to the victim's resource, the attack can force the victim to significantly downgrade its service performance or even stop delivering the service. As compared with conventional DoS attacks, DDoS attacks are more complex and harder to prevent. Confidentiality, Integrity and Availability are core elements of computer security yet human element is the center of all who either unintentionally or intentionally ignore the security requirements which lead to encompass more security break-ins and various vulnerabilities which results to loss of reputation, competitive advantages and money to the organization, which arises the need to work in the area of Web mining security against DDoS attacks. The most common DDoS attacks target the computer network's bandwidth or connectivity [4] that results in degraded productivity. Majorly two attacking techniques i.e. Network based Attacks and Host based Attacks are used for the purpose of DDoS [5]. As they can disable, disrupt and downgrade service performance by exhausting resources. So, it is required to concentrate on the security against such attacks.

This catastrophe may be avoided by implementing security measures [6] & capabilities like Intrusion Detection and Prevention systems to be built into the environment. Keeping this requirement as the only motto, this paper aims to detect the threats of the internet utilities which may cause harm to the users of the internet and finally introduces an anti-phishing mechanism. This mechanism is useful to check the authentication of each website, detection of fake website and prevent the economic damage to the users. The major target of this attack is Net banking sites. Hence, a mechanism called "Anti-phishing System" is developed to perform webpage authentication and webpage detection. This paper comprises of a new anti-phishing algorithm for the end users, which we call Data Shield algorithm. Data Shield algorithm not only detects the known phishing attacks but also the unknown attacks. This developed algorithm is light weighted anti-phishing algorithm and can detect and prevent phishing attacks in real-time.

2. Overview of Phishing Attacks

Phishing is a technique of trying to obtain confidential information such as usernames, passwords and credit card details by impersonating as a trustworthy entity in an electronic communication. Communications claiming to be from popular social web sites, auction sites, web authorities, online payment processors or IT administrators are commonly used to attract the innocent public [7]. Phishing is generally supported out by e-mailspoofing or instant messaging and it often guides the users to enter their particulars at a fake website whose appearance and texture are almost similar to the legitimate sites. Phishing is the best example of social engineering techniques used to cheat users. The recent study showed that out of ten only two of the anti-phishing tools were able to correctly identify over 90% of phishing websites [8]. Therefore, it is important to find effective ways to train people to identify and avoid phishing web sites. They must be taught how to identify phishing URLs, where to look for cues in web browsers, and how to use search engines to find legitimate sites. To distinguish phishing URLs from legitimate ones, they organized URLs into three categories: IP-based phishing URLs, long URLs (with sub-domains), and similar and deceptive domains. However, there are steps that can be taken now to reduce the consumer's vulnerability to phishing attacks [9]. Web-based training materials, contextual training, and embedded training have been shown to improve users' ability to avoid phishing attacks and also give some tips:

1. Don't forget about the URLs
2. Look at the text between the http:// and the first /.The text before the first / (this might be with a .com or .org) is the main domain name.

3. Type the domain name or the organization name into Google search engine. The top result is usually legitimate website.
4. Attackers register domains similar to real sites. Designs and logos can be spoofed. They copy logos and contents from real sites to draw the attention. They request sensitive information. They point all links to real sites to deceive you.

Therefore at present, there is no comprehensive solution in the market to detect phishing. The processes of employing heavy resources to detect emails and to authenticate suspicious web pages are done manually. To avoid such phishing attacks the requirement of an anti-phishing mechanism is major demand.

3. Problem Statement

The entire Internet community is familiar with spam attacks [10]. Phishing is a newer, network based attack but one that results in severe privacy and security violations. Phishing can have the serious problem with the financial ramifications for the individuals and organizations that are targeted. At present, the customer who generally performs E-banking for all his transactions is the prominent victim of phishing attacks. This is because a bank website is more susceptible to frauds as it is full of catchy advertisement for users which resembles the original one [11]. Some of them advertises attractive schemes and many more tricks only to deceive user which is only a hoax. And the users generally believe those ads and provide their credential information on those fake websites. This becomes an advantage for the hackers to use the fraud and scam ads as a medium to access the personal or financial information so they can easily commit cybercrimes. Phishing has become a major concern for ISPs [11], with pressure coming from both users and from the financial institutions targeted by these attacks. But, the lack of public agreement persists for how an ISP should best attempt this. Several techniques [12] have been developed that are currently in use to filter spam. These include: IP address blacklists, Bayesian content filters, content heuristics engines, and content fingerprinting schemes. Although these techniques are effective to varying degrees against spam, only some perform well against phishing. At present no comprehensive solution in the market to detect phishing. Schemes available to authenticate suspicious web pages, work manually. Therefore after studying cons of existing systems, a new anti-phishing mechanism must be introduced which can fulfill the following Functional Requirements of the users

1) To prevent access of phishing web sites
2) To protect important email messages from phishers or attackers
3) To detect fraudulent web site using:

a) Proper DNS and IP matching

b) Authorization of website

c) No manual detection of site is performed; all these functionally should be done automatically.

Therefore, the need for such a security mechanism against phishing attacks give natal to this “Anti-phishing System” which provide an instant, automatic, comprehensive system level solution to perform webpage authentication and Domain name protection against phishing. This mechanism is web utility which is proactively used to detect fraud web sites among the user-defined list of protected web sites. This tool proves as an effective weapon to fight against phishing attacks and provides the best solution.

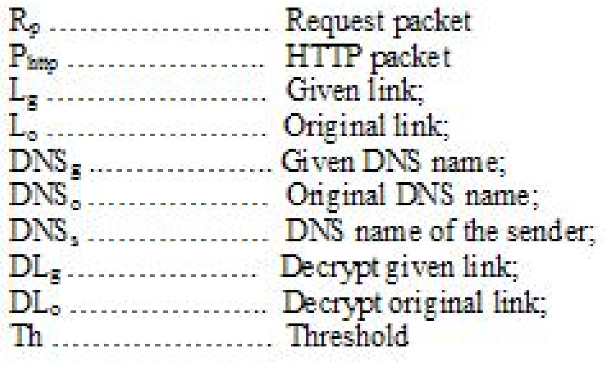
4. Proposed Work

This paper presents a novel approach to analyze a Network based DDoS attack known as Phishing attacks. Phishing is a new type of network based DDoS attack where the attacker generates a copy of an existing Web page to deceive users. They believe that the data is required by their ISP Website and submit their personal, ûnancial data but that is not true. In this paper, the implementation of the proposed anti-phishing algorithm is employed and an anti-phishing mechanism is introduced to avoid such type of phishing attacks and the results are drawn for the same. This proposed anti-phishing algorithm for the end users is named as **“DataShield Algorithm” (DSA)**, which operates on the general appearances of the hyperlinks used in phishing attacks. These appearances and characteristics are derived by examining the documentation of phishing data provided by the Anti-Phishing Working

Group (APWG). This DataShield Algorithm (DSA) proved efficient enough to detect and prevent the known phishing attacks but also the unknown attacks. The purpose of implementing this algorithm is to show that fake links are not allowed to open and if so, then through a proper channel of detection and prevention and the user must get a security alert for the same.

Although, there are many technical or non-technical approaches to stop phishing attacks, to educate users to know how phishing attacks work and be aware of it. This paper focuses on such technical methods to halt the attackers and thus DataShield Algorithm (DSA) is implemented. The purpose of implementing this "anti-phishing mechanism" is to detect and prevent the Network Based DDoS phishing attack whenever the browser collects the HTTP requests from network or hosts. The developed "Anti-phishing System" provides a visibility to get an instant, automatic, comprehensive system level solution to perform webpage authentication and webpage detection against phishing. It gets a proper alert message at the time of entering into phishing site which helps in detecting and blocking the attack in the real time. The proposed DataShield algorithm is given below. Before understanding the algorithm, it is required to be familiar with the parameters of the algorithm.

4.1 The Description of parameters in DataShield Algorithm



R_p	Request packet
P_{http}	HTTP packet
L_g	Given link;
L_o	Original link;
DNS_g	Given DNS name;
DNS_o	Original DNS name;
DNS_s	DNS name of the sender;
DL_g	Decrypt given link;
DL_o	Decrypt original link;
Th	Threshold

Figure1: List of parameters used in Data Sheild Algorithm

4.2 The DataShield Algorithm (DSA) : The Proposed Algorithm

```
/* A request packet containing a DNS name is received from network */
    1.  if ($R_p$! = $P_{http}$)
    2.  return DataShield($L_g$,$L_o$)
    /* Now the DNs request is match in the DataShield Algorithm against 5 types of
Phishing Attacks */
    intDataShield($L_g$,$L_o$)
{
     3. $DNS_g$ = GetDNSName ($L_g$);
     4. $DNS_o$ = GetDNSName($L_o$);
     5. if (($DNS_g$ and $DNS_o$!= Null) && ($DNS_g$! = $DNS_o$))
     6. return Phishing Attacks;
     7. if ($DNS_o$ is dotted decimal IP Address)
     8. return Likely Phishing Attacks;
     9. if($L_o$ or $L_g$ is coded)
{
     10. $DL_g$ = decrypt($L_g$);
     11. $DL_o$ = decrypt($L_o$);
     12 return DataShield($DL_g$, $DL_o$);
}
      /* analyze the domain name for Likely Phishing Attacks */
     13. if($DNS_g$ != any DNS name or dotted IP address)
     14. returnDNSAnalyzer($L_o$);
}
intDNSAnalyzer($L_o$)
{    /* it uses its clean and unclean list to analyze the original DNS name */
     15.  if ($DNS_o$ is in cleanlist)
     16.  return Not a Phishing Attacks;
     17. if ($DNS_o$ is in uncleanlist)
     18. return Phishing Attacks;
     19. returnPatternSimilarity($L_o$);
}
     intPatternSimilarity($L_o$)
{
     20. if ($DNS_s$ and $DNS_o$ are not similar)
     21. return Likely Phishing Attacks;
     22. for (each earlier DNS$_g$ in the cookie)
{
     23. A = Likeness (earlier DNS$_g$, L$_o$);
     24. if(A = = true)
     25. return Likely Phishing Attacks;
}
     26. return Not a Phishing Attacks;
}
float Likeness (seq, L$_o$)
     27. if (seq is portion of L$_o$)
     28. return true;
     29. intmax_len = the maximum length of sequence and DNS$_o$;
     30. intmin_alter = the minimum no. of alterations to transform seq to DNS$_o$(0r
vice-versa);
     31. if (Th<(max_len-min_alter)/max_len<1)
     32. return true;
     33. return false;   /* Thus, the user receives the request in proper URI format */
```

4.3 Functioning of the proposed DataShield Algorithm (DSA)

The DataShield Algorithm (DSA) aims to detect fraudulent web site using Proper DNS matching and authorization of websites. It also compare between original link and the given link. It also calculates the Likenesses of a URI with a known trusted site. The DataShield algorithm (DSA) checks for five types of possible phishing attacks. Type I: If given DNS and original DNS are different. Type II: If DNS contains dotted decimal. Type III: If DNS is not found in the clean list. Type IV: If DNS of the sender is different from original DNS. Type V: If given DNS is similar but not identical. In all these above cases, there is large possibility of the occurrence of phishing attacks. The algorithm works as follows. After extracting the user request, if it is not in HTTP form it is send to the DataShield module as the request packet is a DNS name. This module matches between the given and the original DNS names, if these names are not similar, then it is a phishing attack. If the original DNS directly contains the dotted decimal IP address, then it is a likely phishing attack. In case of the original link or the given link is encrypted then the links are ûrst being decoded. In case, if given link contain no information about the destination the algorithm calls DNSAnalyzer module to examine the original DNS. This module of DataShield Algorithm (DSA) checks the availability of original DNS in its two available lists i.e. clean and unclean list and decide accordingly as a phishing attack or not. If it is not found in both these lists then the original DNSwill call the PatternSimilarity module. This module is used when the clean and unclean list fails. It is used to handle unknown attacks. PatternSimilarity module compares the original DNS name of a hyperlink and DNS name of sender, and found it is similar with one or more name in the cookie and invoke the Likeness module otherwise declare it as the attack. Likeness calculates the likelihood of original DNS and the earlier DNS present in the cookie by taking ratios. If the likeliness threshold is found satisfactory then user request is acknowledged or fulfilled otherwise a phishing attack is being displayed.

5. Performance Analysis of Anti-Phishing Mechanism:

Due to the current dangers occurred because of phishing attacks from home to offices. There is an essential demand to develop such an application which can work for the security against phishing attacks. For this purpose, this application is developed and implemented on the Standalone side which always requires a net connection for the interaction of links taken in the application. User's click can take him to another page which may or may not be safe. The user must be aware of all fake links & should not be allowed as this may be an attempt of phishing attacks and if they occur, they provide a security alert which warns the user against fake site which may register him to phishing attacks. In this way, the attacks are detected and the particular site or link can be blocked to prevent the user from phishing attacks. Similarly, this anti-phishing mechanism can be applied on the other links and can work efficiently in the real time.

The below is the flowchart of used Anti-Phishing Mechanism in Figure 5.2. This flowchart helps in understanding the flow of a data request in this proposed anti-phishing mechanism.

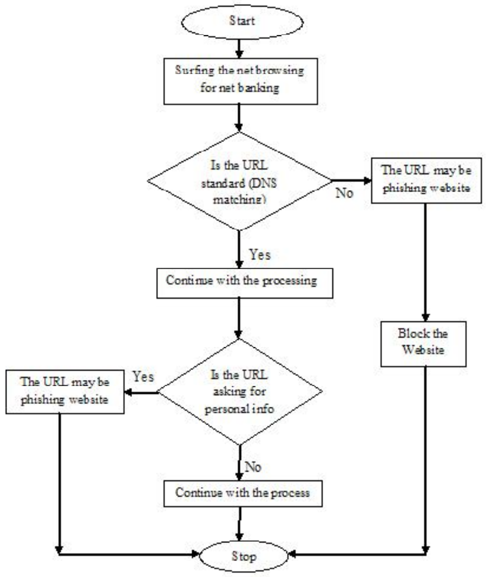

Start
Surfing the net browsing for net banking
Is the URL standard (DNS matching)
No
The URL may be phishing website
Yes
Continue with the processing
Block the Website
Yes
The URL may be phishing website
Is the URL asking for personal info
No
Continue with the process
Stop

Figure 2: Flowchart of used Anti-Phishing Mechanism

6. Conclusion

Phishing is a novel kind of network attack where the attacker creates a duplication of an accessible Web page to deceive users into submitting private, monetary, or key data which not only distress the network security but also lead to large monetary loss from home to corporate offices. All the existing system are not sophisticated enough to reach a satisfactory security level. Therefore, this paper implemented a new Anti-Phishing mechanism that improves the security against Phishing Attacks. The paper also proposed an anti-phishing algorithm for the end users named as “DataShield Algorithm” (DSA), which operates on the general appearances of the hyperlinks used in phishing attacks. These appearances and characteristics are derived by examining the documentation of phishing data provided by the Anti-Phishing Working Group (APWG). After the implementation of DataShield Algorithm (DSA) it is proved that it is efficient enough to detect and prevent both the known and unknown phishing attacks. It is also capable of Domain name protection and Webpage appearance protection.This developed algorithm is light weighted anti-phishing algorithm and can detect and prevent phishing attacks in real-time. We can use this anti-phishing system to prevent from phishers and their disastrous attacks in the real world. In future, the possibility of developing a secure Hybrid Intrusion Detection System is searched out.

References

1. Shrivastava, G., Sharma, K. and Rai, S., "The Detection & Defense of DoS&DDos Attack: A Technical Overview" IEEE International Conference on Computer Engineering and Technology, 2010.
2. Jangra, A. and Verma, B., "Web Mining Security: An Overview" in International Journal on Advance Computer Technology, Vol. 1, 2011.
3. "Survey of Network-based Defense Mechanisms Countering the DoS and DDoS Problems" ACM Transactions on Computational Logic, Vol. 2, No. 3, 2006.
4. Jain, P., Jain, J. and Gupta, Z. "Mitigation of Denial of service (DoS) attacks" in International journal of Computational Engineering & Management, Vol. 11, Jan 2011.
5. Paul J. Criscuolo. "Distributed Denial of Service Trin00, Tribe Flood Network, Tribe Flood Network 2000, AndStacheldraht CIAC-2319". Department of Energy Computer Incident Advisory Capability (CIAC), February 14, 2000.
6. JelanaMirkovic, Janice Martin, and Peter Reiher, "A Taxonomy of DDoS Attacks and DDoS Defense Mechanisms", UC Technical Report, 2002.
7. KirdaEngin and Christopher Kruegel, "Protecting Users Against Phishing Attacks with AntiPhish", Technical University of Vienna, 2006.
8. Zhang, Y., S. Egelman, L. Cranor, and J. Hong, " Phinding Phish: Evaluating Anti-Phishing Tools" in Proceedings of 14th Annual Network and Distributed System Security Symposium (NDSS 2007), San Diego, CA, 28 Feb -2 Mar, 2007.

9. Kumaraguru, P., Y. Rhee, A. Acquisti, L. Cranor, J. Hong and E. Nunge. 2007. "Protecting People from Phishing: The Design and Evaluation of an Embedded Training Email System" in Proceedings of Computer Human Interaction, CHI 2007.

10. Gregg Tally, Roshan Thomas and Tom Van Vleck, "Anti-Phishing: Best Practices for Institutions and Consumers" McAfee Research Technical Report, September 2004.

11. "Anti-Phishing Best Practices for ISPs and Mailbox Providers" jointly produced by the Messaging Anti-Abuse Working Group (MAAWG) and the Anti-Phishing Working Group (APWG), Version 1.01, July 2006.